Ontologies for Multimedia Objects and Workbench. Final Version.

SALERO Deliverable 3.1.7

IST FP6-027122

Ontologies for Multimedia Objects and Workbench. Final Version.

SALERO Deliverable D3.1.7

SALERO identifier:	SALERO-D317-JRS-Ontologies-and-workbench-v07.doc
Deliverable number:	D3.1.7
Author(s) and company:	G. Kienast, W. Weiss, P. Hofmair (JRS); T. Bürger, S. Huber (LFUI); P. Swamy, R. Villa (UG);
Work package / task:	WP03
Document status:	Final
Confidentiality:	Public

Version	Date	Reason of change
1	2009-08-07	Document created. Included Evaluation report. Call for comments & input
2	2009-08-31	Added input on evaluation (User Group feedback)
3	2009-09-09	Added input on annotation tool , ontologies & evaluation
4, 5	2009-09-22	LFUI input on ranking & evaluation
6	2009-09-29	Version for internal review (to be done by DIT)
7	2009-10-08	Final version to be submitted to the Commission

The work presented in this document was partially supported by the European Community under the Information Society Technologies (IST) priority of the 6th framework programme for R&D.

Table of Contents

1 Executive Summary

This deliverable is the seventh and final deliverable of Work Package 03 ("Media Semantics and Ontologies") and presents the achievements during the final phase of the SALERO project. In SALERO WP03 main partners JRS and LFUI have jointly developed a set of ontologies and tools designed for semantic annotation and search with the overall goal of allowing easy reuse of assets in digital media production.

The *ontologies* developed in SALERO are described in Chapter 3, which starts with a summary of the user partners' requirements and the ontology engineering paradigm.

Chapter 4 contains a detailed description of the *Semantic Workbench* and the *semantic services*.

Chapter 5 presents the *Intelligent Media Annotation and Search* (IMAS) Tool based on the semantic services. The tool uses an approach called statement-based annotation. The design and functionality of tools for Semantic Annotation and Search are described and the concept of statement based annotation is introduced. In section 5.4 semantic-based resource ranking is discussed.

Chapter 6 is dedicated to *evaluation*, *user group feedback* and *SWOT analysis* and *measures taken* based on evaluation.

2 Introduction

2.1 Scope and Purpose of this Document

This is the last of a series of deliverables in Work Package 03. Based on previous deliverables it describes the final versions of the ontologies and the semantic tools and services developed in SALERO. It also describes evaluation efforts and feedback received from the SALERO user group.

Wherever possible and applicable this deliverable references previous documents produced by this (and other) work packages to avoid duplicate description of already documented work.

2.2 Status of this Document

This is the final version of D3.1.7.

2.3 Related Documents

More information can be found at the SALERO website which can be found at http://www.salero.eu.

It is also recommended to read the following Deliverables

- D3.1.1 Representation Techniques For Multimedia Objects
- D3.1.3, D3.1.5: Ontologies for Multimedia Objects and Workbench
- D2.3.4: Updated User Requirement Document
- D5.5.4: The Design and Implementation of the Context-based Search System

3 The SALERO Ontology

This section gives a short summary on the requirements for ontologies in SALERO, the process taken and the ontology itself.

The first two sections are kept short, because requirements have been described in Deliverable 2.3.4 and the ontology engineering process was documented in Deliverable 3.1.5. The Virtual Character Ontology is described in Deliverable D3.1.5 and [Bürger et al. 2008]. The SALERO Asset Retrieval and Media Relationship ontologies are described in Deliverable D3.1.6. For details, please refer to these documents.

3.1 Requirements

This section gives a brief overview on the application scenarios of the user partners in SALERO. For a full description please refer to D3.2.4 “Updated User Requirements”.

The core domain shared by the user partners is the specification and description of virtual characters to enable retrieval of characters according to their properties. The main expected benefit of the ontologies is to make multimedia assets searchable to allow easy and fast re-use of existing assets. Annotation, search and retrieval of assets is achieved via the Semantic Annotation & Search Tool.

3.1.1 BLITZ

For asset management in the games development pipeline, there are two co-existing areas that could potentially benefit from ontology based metadata. The first case is management of legal and IP rights, the second on more generalised search and retrieval of assets based on their content.

3.1.2 AM / FBM-UPF

The Program Editor is a software for creating 3D animations automatically. A vital functionality is to be able to automatically select the right content in terms of fitting to a character’s personality, mood, emotions, context and situation.

3.1.3 TAIK

The motivation of using ontologies in 3D content production is in the research phase and during character animation for retrieval of inspiring related content before starting content creation or of parts / characters to reuse during a production.

3.1.4 PGP

Assets need to be managed for use in creation of productions in three different domains: video, publishing and interactive. The ontology shall support efficient search & retrieval in this database for fast and easy asset retrieval and use.

3.2 Ontology Engineering

SALERO has followed an ontology engineering approach based on work described in ([Uschold et al, 1995], [Grüninger et al, 1995], [Fernández-López et al, 1999], [Staab et al 2001] and [Noy et al, 2001]. The following steps were performed: Determine the Domain and Scope of the Ontology

1. Determine Domain and Scope of the Ontology
2. Consider Reusing Existing Ontologies
3. Enumerate important terms in the ontology
4. Define the Classes and the Class Hierarchy
5. Define the Properties of Classes

6. Define the Restrictions of the Properties
7. Create Instances

Above mentioned steps are described in more detail in D3.1.4.

The Virtual Human Ontology developed by the aim@shape project has been identified as a well-suited starting point for ontology development in SALERO. Based on this the SALERO Virtual Human Ontology was developed.

3.3 Ontology Description

To represent the semantic annotations we developed two ontologies according our needs. The first one is the SALERO annotation ontology (c.f. Figure 1), for describing media resources, authors of media resources, projects in which they are used, and annotations. The property annotea:related, derived from the Annotea Annotation Schema[1], is used to describe the annotation statements. The subclasses of Annotation and AnnotationConcept form the domain specific parts of the ontology, which have to be adapted if someone wants to use the annotation tool in a different domain. The subclasses of Annotation describe media specific annotation types such as annotations of audio, image or video.

The subclasses of AnnotationConcept represent the domain-specific part of the annotation ontologies and are used to describe the content of media resources. They currently include the following classes:

- **Character:** The actors of a movie, e.g. Bing and Bong.
- **CharacterParts:** Parts of the body of the characters, e.g. Hand, Nose, or Toe.
- **Expression:** Includes verbs and adjectives to describe the behaviour of characters, e.g. smiling, dancing, open, or wet.
- **Location:** A geographical or imaginary location, e.g. Barcelona or Outer Space.
- **Object:** A tangible or visible entity, e.g. Balloon, Umbrella, or Cake.

[1] Annotea Annotation Schema http://www.w3.org/2000/10/annotation-ns

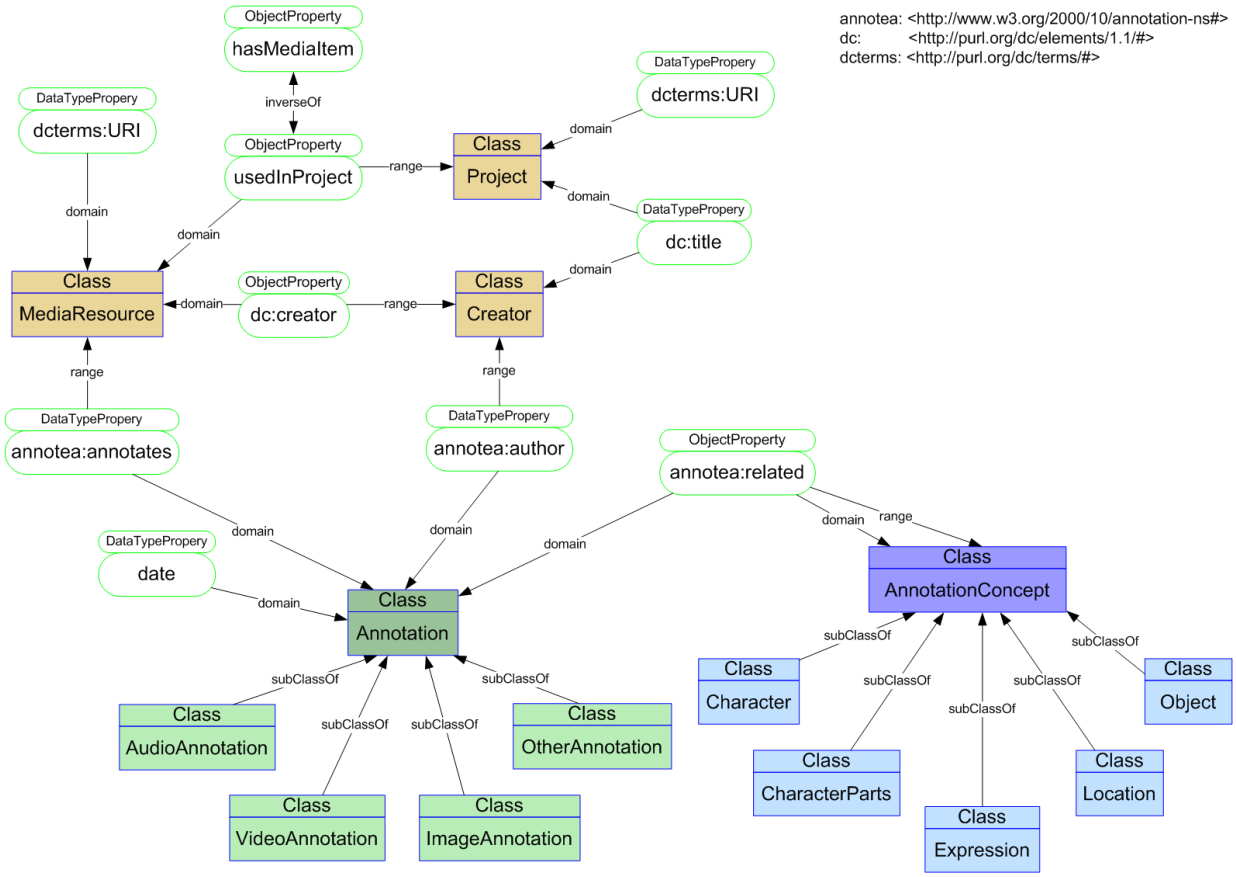

Figure 1: The SALERO annotation ontology

The scope of the second ontology is to describe relationships between media resources and how they are most probably derived or based on each other. The relationships of the Functional Requirements for Bibliographic Records (FRBR)[2] model provide a solid ground to describe the relationships of media resources in the media production domain in general and in the SALERO project in particular. The relationships are supposed to enhance the browsing experience of media professionals in their media collections.

The relationships of our ontology were chosen based on a general analysis of the domain based on related work, an analysis of images from the MyTinyPlanets[3] collection, and a set of expert interviews. It contains a subset of the FRBR core ontology[4] and two additional relationships, which were not covered by the FRBR core ontology (frbr: http://purl.org/vocab/frbr/core#) (c.f. [FRBR, 2009, Riva 2007]):

- **frbr:adaptation** A media resource, which is based on another media resource, exchanges parts of the original content, e.g. in a new image a tree is exchanged by a space ship.
- **frbr:alternate** An alternative (file) format of the same content, e.g. jpg instead of png.
- **frbr:imitation** A real world scene is imitated in a cartoon, e.g. a scene of “Star Wars" is imitated by Bing & Bong in “Tiny Planets".
- **frbr:part** A media resource contains a part of another media resource, e.g. the hands of the character Bing are a part of the final animation of Bing. Supposing that the hands are modelled in a different file.

[2] Functional Requirements for Bibliographic Records (FRBR) http://www.ifla.org/VII/s13/wgfrbr/index.htm

[3] MyTinyPlanets http://www.mytinyplanets.com/

[4] FRBR Core ontology http://vocab.org/frbr/core.html

- **frbr:reconfiguration** A rearrangement of the content of a media resource, e.g. a new scene is based on an existing scene with the same content such as trees, space ships, characters. In the new scene the content is locally rearranged.
- **frbr:revision** A newer or other version of a media resource, e.g. a new version of a space ship, which is based on an existing space ship, uses a different texture for its surface.
- **frbr:transformation** From a sketch to a (3-d) model, e.g. the first step in creating a 3-d model is to draw a sketch on a sheet of paper. The next step is to create a 3-d model in the computer, which is a transformation of the sketch.
- **frbr:translation** A translation into a different language, e.g. the embedded text and audio of a clip are translated from English to French.
- **duplicate** For duplicated media resources, e.g. the same file is stored at a different location.
- **version** The super property for frbr:adaption, frbr:reconfiguraiton, frbr:revision and frbr:transformation.

The property **relationships** is the super property for all relationship properties, its range and domain is MediaResource.

A detailed description of the SALERO annotation and relationship ontologies can be found in the deliverable 3.1.6 [Bürger et al., 2009].

4 Semantic Workbench and Services

In order to pave the way for the use of ontologies and semantic technologies in media production, SALERO developed a management framework for multimedia ontologies, tools to annotate existing media resources and semantic search facilities to retrieve resources based on semantic annotations. The framework, which offers these functionalities, is introduced in the following.

The *SALERO Semantic Workbench* supports the life cycle of multimedia ontologies, that is their creation, management, and use. This includes the following main functionalities (cf. Figure 2):

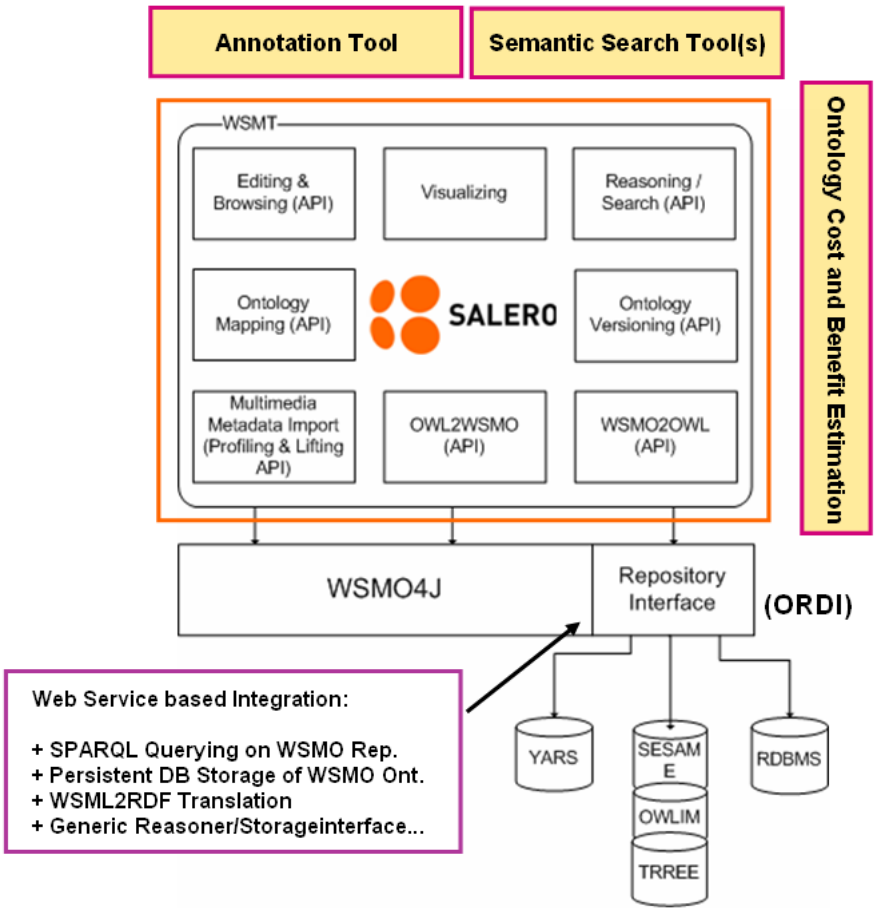

Figure 2: The SALERO Semantic Workbench (High Level Architecture)

- **Ontology Management** whose central aspects include manual and semi-automatic creation of domain ontologies, alignment of different domain descriptions, translations of ontologies, versioning, or storage of ontologies.
- **Annotation Support** whose central aspects includes the support for non-technological users with the annotation of media items which is realized in several annotation tools.
- **Semantic Search Support** which offers advanced retrieval capabilities based on semantic annotations.

To realize this functionality, the workbench not only offers a graphical user interface to engineer ontologies, but also a set of services which provide ontology management functionality to other applications. The workbench acts in the background and its central aspects are thus realized as an API which is designed with the aim to integrate the functionality needed for semantic media annotation and semantic search into plug-ins and interfaces of other applications. This includes most notably storage, querying, or retrieval of annotations.

The functionalities offered by the workbench and its services are briefly sketched in the following. More details about the requirements posed to it and parts of its functionality are described in [Bürger et al. 2007] and [Kienast et al. 2008] respectively. We particularly detail functionality which has not been previously reported.

4.1 Ontology Management Support

The part of the workbench for the management of ontologies is based on the Web Service Modeling Toolkit (WSMT)[5] that among others provides a set of graphical tools for the engineering of WSMO ontologies, for the interaction with external tools such as execution environments and repositories [Kerrigan et al. 2007]. WSMT is a collection of tools for the engineering of Semantic Web Services and ontologies implemented in the Eclipse framework.[6] In SALERO we added the possibility to persistently store and access ontologies in an ontology repository as realized by the **Repository Service** described below.

Engineering ontologies is most notably supported by functionalities offered by WSMT which has been described in D3.1.5 [Kienast et al. 2008]. A screenshot of the recently released WSMT 2.0 is presented in Figure 3:

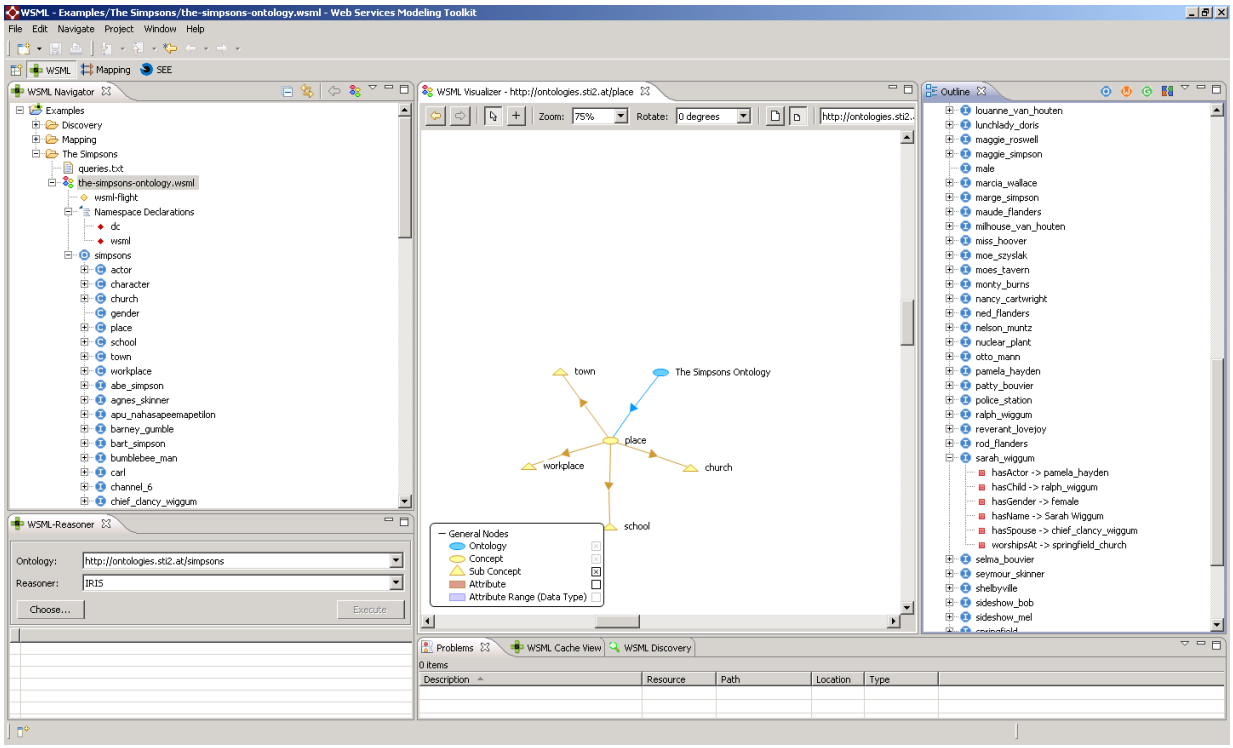

Figure 3: WSMT 2.0

4.2 Ontology Life Cycle Support: The Workbench Services

The services offered by the semantic workbench include:

- **The Repository Service** which offers an API for the persistent storage of WSML[7] ontologies and their elements (e.g., concepts, properties, axioms). It supports management of these elements and the execution of SPARQL queries. The service is realized on top of the **Ontology Representation and Data Integration (ORDI)** -- framework, which most notably provides a

[5] http://sourceforge.net/projects/wsmt/

[6] http://www.eclipse.org/jdt/

[7] http://www.wsmo.org/wsml/wsml-syntax

scalable repository implementation, a WSMO-RDF parser, serializer, and access to query and reasoning facilities.[8]

- **The Annotation Service** is concerned with the management of semantic annotations and provides an API to manage and validate annotations against the ontologies stored in the repository.
- **The Semantic Search Service** offers an API to search for ontology elements and additionally offers keyword-based search for annotations which is expanded into full-text queries on a generated index and SPARQL queries. The service supports two different types of queries: The first option are concept-based queries which are translated into SPARQL queries and which are evaluated via the **Repository Service**. The second option are keyword-based queries which are executed in a Lucene[9] index which is generated based on the ontologies and instance data in the repository.
- **The Ranking Service** offers functionality to rank media resources based on semantic annotations. This service is used by the semantic search and the recommendation service. The ranking functionality is described in more detail in Section 5.4 below.
- **The Recommendation Service** offers an API for retrieval of ontology elements which are prominently used for annotation and gives recommendations of related results during search. This in turn is used for ranking in search but also for ranking the elements which are retrieved via the *Repository Service*.
- **The Profiling and Lifting Service** can be used to extract structural semantic information from existing MPEG-7[10] documents and for their semantic enrichment. This service is described in more detail in [Bürger et al. 2006].

The APIs for using the services and their functionalities are described in more detail in D3.1.5 [Kienast et al. 2008].

4.3 Ontolgy Use and Evolution: The Intelligent Media Annotation and Search System

Besides tools for ontology management, the workbench is accompanied with a tool set for annotation and semantic search: the **SALERO Intelligent Media Annotation & Search**} (IMAS) system [Weiss et al. 2009]. The IMAS integrates annotation and search into one application and further provides access to content-based search facilities. Both semantic search and content-based search can be accessed via a single interface and the results are being fused. In order to adequately support non-experienced users in annotation, theworkbench offers a methodology to support the users including (i) the selection of adequate ontology elements and (ii) the extensionof ontologies during annotation time and by that supports ontology use as well as evolution (cf. [Buerger and Ammendola 2008]).

The IMAS system is discussed in more detail in Section 5.

Ontology evolution is supported by the **Element Selection and Element Addition (SA) Methodology**, which is a manual annotation methodology that allows non-experienced users to annotate multimedia content using elements from ontologies in an easy and abstrusive way. The methodology combines several semantic-based approaches from the information retrieval domain (e. g. [2] or [3]) and recent proposals for end user-driven semantic content creation [1, 4] in one coherent framework.

The aim of the SA methodology is to support users in the process of manual annotation by providing aid in the selection of adequate ontology elements and in the extension of ontologies during annotation time. The methodology is designed to be integrated in media production environments to be available just in time of the creation of new media elements.

[8] http://www.ontotext.com/ordi

[9] http://lucene.apache.org

[10] http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm

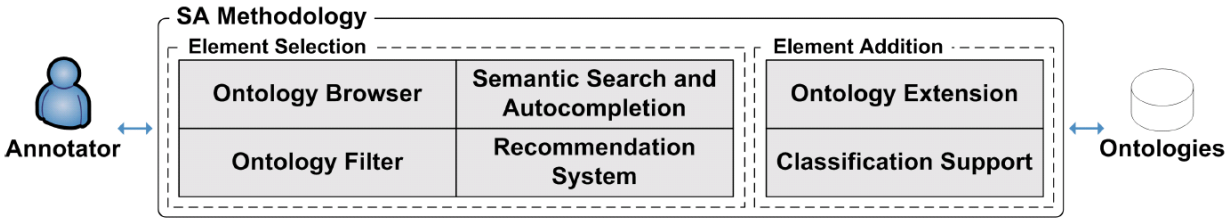

Figure 4: The SA Methdology (cf. [Bürger and Ammendola 2008])

Figure 4 depicts the structure of the methodology which consists of two parts:

- The **Element Selection** - part includes a set of components for the selection of ontology elements based on semantic techniques presented in [2] and [3]. The key idea is to help the user in finding adequate ontology elements by using semantic retrieval techniques and by automatically proposing ontology elements that could be relevant for annotating a specific resource.
- The **Element Addition** - part consists of two components enabling a collaborative and work-embedded ontology engineering approach based on mechanisms presented in [1] and [3], and the classification mechanism used in [4]. These components provide the possibility to add missing elements to ontologies while working with them, to adapt previously added elements and to provide support for classification through a manual and a semi-automatic classification approach. For the addition of elements, the user is guided through the existing elements of the ontologies and for the classification; elements are proposed based on a recommendation system.

5 Semantic Annotation & Search

The management of media resources in media production is a continuous challenge due to growing amounts of content. Because of the well-known limitations, manual annotation of media is still required. We present a statement-based semantic annotation approach which allows fast and easy creation of semantic annotations of media resources. The approach is implemented in the Intelligent Media Annotation & Search[11] (IMAS) system. An integral part of the work being done in SALERO is the management of media objects with semantic technologies which is addressed by the IMAS system by enabling their semantic annotation and retrieval. The use of semantic technologies reduces the problem of ambiguity in search by using existing, well-defined vocabularies; it allows us to do query expansions, and to deal with multiple languages.

During prototypical development iterations of our system we have experienced, that most paradigms applied in semantic annotation tools are not suitable for inexperienced users who are typically used to keyword-based tagging and suffer from information overload when confronted with complex annotation tasks and user interfaces. Our aim was thus to develop an approach which is faster and easier to use for our targeted user group, while making a compromise in complexity of full semantic annotations. Beside describing the content of each media resource, the approach allows to relate media resources to other media resources. In the following we present the IMAS system and the implemented annotation approach. The system is based on the principles described in [Bürger and Ammendola, 2008]. These principles describe methodologies to support users in the process of manual semantic annotation including (i) selection of adequate ontology elements and (ii) extending of ontologies during annotation time. Furthermore, we present an evaluation of different user interface techniques for creating annotations.

5.1 System Description & Design Principles

The main aim of the IMAS is to allow easy annotation of media assets for later retrieval and reuse by users in media production. In order to support this, it has been built based on the following design principles:

1. **Designed for content creators.** The target users of the system are nontechnical experienced content creators in the domain of media production.
2. **Easy to use.** The interface provides Web 2.0 - based interaction mechanisms to make the annotation process as easy as possible.
3. **Global annotations.** To facilitate the annotation process, we only allow global annotation of media resources instead of annotating parts of it.
4. **Statement-based annotation process.** We allow to create statements, which use ontological elements, to describe the content of media resources.
5. **Ontology extension during use.** We allow users to easily extend the ontology during use based on principles described in [Bürger and Ammendola, 2008].
6. **Portability of the system.** In order to port the systems to other domains, only the underlying annotation ontology has to be adapted.
7. **Integration of semantic and content-based search.** The system provides an integrative view onto results from different search engines and by that provides a fallback solution which is able to retrieve objects without annotations too.

The IMAS integrates two systems whose functionalities are offered as a set of Web services, i.e. the Semantic Services and the Content-based Services. The architecture of the IMAS system is shown in Figure 5:

[11] IMAS Online Prototype: http://salero.joanneum.at/imas/

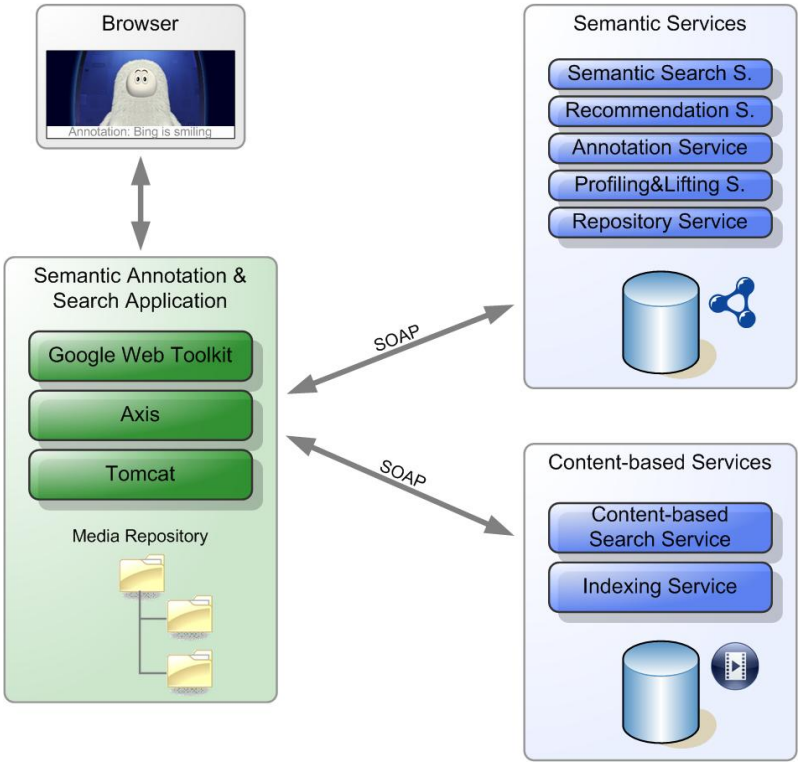

Figure 5: IMAS System Overview

5.1.1 Services

The Semantic Services are described in section 4.

The Content-based Services offer functionality for the indexing and retrieval of image, video and textual information. The aim of the service is to complement the semantic search services and as a fall-back system in the cases where material is not indexed by the semantic services. As such their emphasis is on automatic indexing techniques, which can be used to retrieve images, text or video without manual annotation. Textual information is indexed using standard Information Retrieval techniques (the Terrier system is used [Ounis et al., 2006]); image and video data is indexed by extracting low-level visual features based on the MPEG-7 standard, as currently implemented in the ACE toolbox [O'Connor et al., 2005].

5.2 Statement-based Annotation

The IMAS end user application is an integrated Web-based application which can be used to annotate and search for media objects. As illustrated in Figure 5, it consumes functionality of (i) the Semantic Services which are used to add, update and delete semantic annotations as well as to search the annotation repository and (ii) the Content-based Services which are used to retrieve media resources based on its intrinsic features such as colour, histograms or shapes.

Figure 6: IMAS Screenshot of Annotation view.

5.2.1 Usage

The application allows to annotate arbitrary resources which are stored in pre-configurable media repositories. In order to ease the annotation process for our target user group, media resources are annotated globally instead of region or segment-based. Media resources are annotated by creating statements and by relating them to other media resources. Annotation statements contain semantic elements which are defined in the annotation ontology (see also section 3). The annotation statements are formalized according to the annotation ontology and represent natural language-like statements about the content of the media resource. Statements are in the form of

$$< Concept \ \mathbf{isRelatedTo} \ \{Concept\ 1 \ldots Concept\ n\} >$$

which are triples where a concept can be set in relation to other concepts. Using statements with semantic elements is a compromise in complexity between loose and fully semantically described annotations. Figure 7 illustrates statements with an example Image from the tiny planets[12] universe.

- Bing is related to: Bong, Alien, reading, book
- Bong is related to: smiling
- Alien is related to: surprised

Figure 7: Example of annotation Statements.

To create such statements, three different input options are available as shown in Figure 8: (1) combining concepts via drag-and-drop, (2) selecting concepts consecutively and (3) using the text box as a command line interface in the spirit of [Raskin, 2008] with auto-completion. Input option three is

[12] My Tiny Planets: http://www.mytinyplanets.com/

optimally suited for frequent users and input options one and two are ideal for users who rarely create annotations.

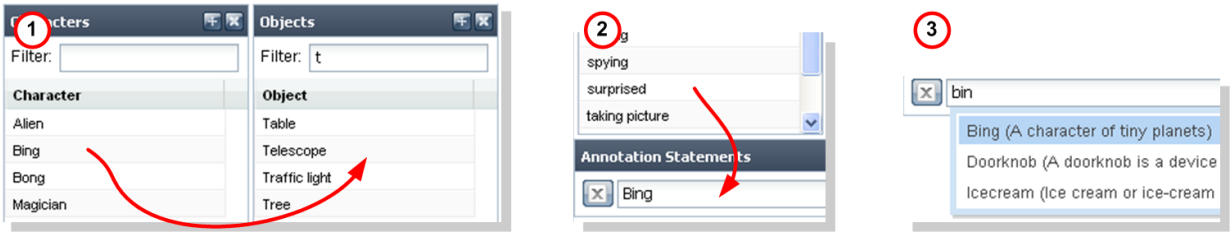

Figure 8: Creation of annotation Statements.

An additional possibility to annotate the content of media resources in the IMAS system is to relate them to each other. Hereby, we can describe that one media resource is, for instance, a part, a revision, or a transformation of another media resource. This allows us to use statements of the related media resources, to keep track of revisions of the media resources or to suggest alternatives in the search result. The relationship ontology with its properties is described subsequently in section 3.3. To create relationships (see also Figure 9) of selected media resources (1), the user drags a source media resource from the file browser (2) and drops it on the desired relationship (e.g. the revision) of the relationship panel (3). This action creates the following annotation (4):

< *SourceResource* ***is a revision of*** *TargetResource* >.

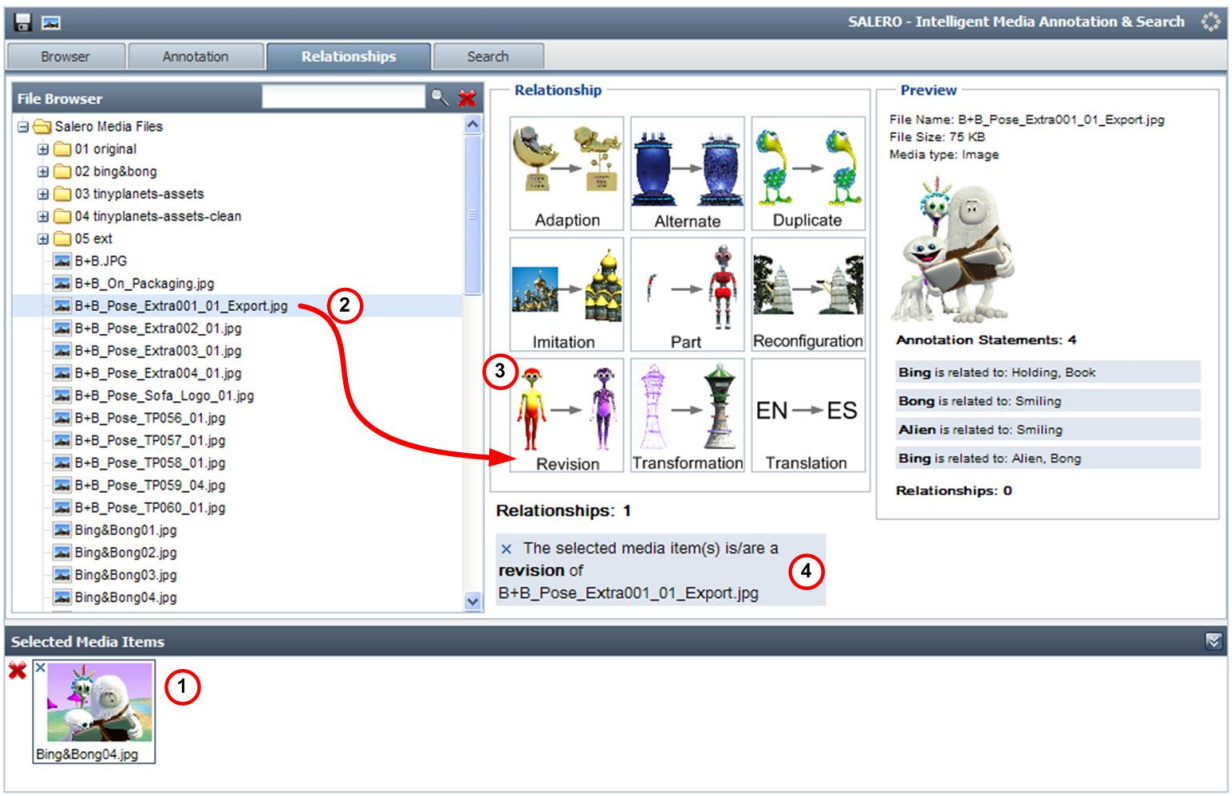

Figure 9: Creation of relationships

5.3 Semantic Search

To search for media objects, the following input options are available to create a query (see also Figure 10): (i) free text, (ii) semantic concepts, (iii) statements, and (iv) images. Free text search is both executed in the Semantic Services and the Content-based Services. The concept-based search is expanded in the Semantic Services. Via the exemplary images a query is submitted to the Content-based Services. The results of both systems are merged based on the respective rank values being returned.

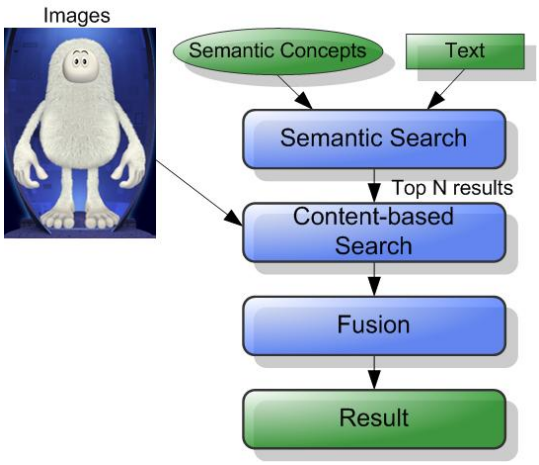

Figure 10: Search Workflow

5.3.1 Semantic Search Algorithm

The Semantic Services support two different types of queries: The first option are concept-based queries which are translated into SPARQL queries and which are evaluated via the Repository Service of the Semantic Services in the repository. The second option are keyword-based queries which are executed in a Lucene[13] index which is generated based on the ontologies and instance data in the repository. The Lucene index preserves the annotation triples and also stores subsumption information to enable efficient retrieval. Keyword and concept queries can be mixed to increase the precision of keyword-based querying in the system. Furthermore selected WordNet[14][15] relations of the concepts can be expanded to increase the result set while preserving precision. Once the queries are answered, the results are ranked based on the semantics of the annotations attached to the returned results and their degree of match of the query. The ranking is based on the popular tf/idf measure which is customized for semantic annotations to take into account triples instead of just terms and the subsumption hierarchies.

5.3.2 Search Result Fusion

The core problem of generating the final list lies in devising a mechanism of combining the results from each system to produce effective results through the IMAS engine. The problem of combining results is widely studied, both by information retrieval communities as well as multimedia retrieval groups at various levels: combining results from many query examples, combining results originating from multiple features or results obtained from different search engines. Multimedia information retrieval communities refer to this as multimodal fusion (cf. [Mc Donald and Smeaton, 2005], [Hoi and Lyu, 2008], [Qi and Han, 2005]) and querying with multiple examples (cf. [Westerveld and Vries, 2004], [Kludas et al., 2008]).

In our system a round robin mechanism combined with a polling-based result fusion mechanism is adopted to fuse results obtained from semantic and content-based search systems. A keyword-based search is employed to obtain the initial list of n results from the semantic search system. The top m << n results are then used by the content-based search to retrieve more visually similar images. Since it is also important to select the best feature or re-order the features according to their preference for the query examples favouring to obtain more diverse results, the knowledge embedded in queries are exploited by computing the correlation and t-distribution in the queries. The feature which has less correlation and high t-distribution is chosen as the most useful feature.

[13] Apache Lucene: http://lucene.apache.org/

[14] WordNet: http://wordnet.princeton.edu/

[15] WordNet Ontology: http://www.w3.org/2001/sw/BestPractices/WNET/wordnet-sw-20040713.html

5.4 Semantic-based Resource Ranking

The semantic-based resource ranking used in the semantic workbench has been implemented in two iterations. The initial ranking followed the idea of the approach presented in (Ruotsalo, et al., 2007). This approach is inspired by the vector space model[16] in which query vectors are compared to document vectors and documents are then ranked according to their similarity with the query. In analogy to the vectors, triples are used instead in this approach. Accordingly, the keywords from the query are transformed into triples and compared with the source triples annotating resources. The initially implemented approach was evaluated and, based on that, refinements have been suggested: During an initial user evaluation various weaknesses could be observed based on which the algorithm has been adapted and re-implemented. The ineffectiveness of initial ranking algorithm has various reasons including the inappropriateness of query vectors constructed and the contradictory influence of weighting factors on the overall ranking score.

5.4.1 Refined Ranking Approach

According to the observed weaknesses of the initial ranking approach, several refinements have been implemented. The refinements are presented in the following, accompanied by a motivating example.

In a search for “Bing umbrella” the right image in Figure 11 is more relevant to the query and should be ranked higher than the left one. In the initial ranking approach this was not achieved and therefore the refinements change this.

Bing related to Holding
Bing related to Ice cream
Bong related to Smiling
Umbrella

Bing related to Waving
Bong related to Smiling
Rock
Bing related to Umbrella

Figure 11: Exemplary Images with Annotations

The **first refinement** changes the way how target triples are computed in order to reflect the user’s information need best.

The **second refinement** substitutes the weighting factor “concept frequency” which accounts for the frequency of concepts appearing in both query and annotations with a factor called “triple boost factor.” The “triple boost factors” accounts for the importance of triples in annotations according to their type.

According to the **third refinement** matching scores for query triples and document triples are computed only once instead of multiple times as it was the case in the initial approach.

The **fourth refinement** introduces a triple hit factor, which is determined by the amount of matching target triples against a resource’s source triples. The triple hit factor is additionally multiplied by a weighting factor depending on the type of matching triples.

[16] http://en.wikipedia.org/wiki/Vector_space_model

5.4.2 Evaluation

In this section the refined ranking approach presented in the previous section is evaluated. We present the experiment performed to evaluate the approach and the experimental results.

Experiment

For the evaluation of the refinements of the ranking algorithm presented in the previous section an end-user experiment was performed. Therefore a panel of 5 probands was chosen to play the role of a ranking algorithm. In 3 runs each proband had to order a set of resources according to their relevance to a query. To perform the experiment mages from Tiny Planets were chosen. The images with their annotations have been put on a piece of paper. An example is shown in Figure 12.

Figure 12: Examplary Image

To perform the experiment no extensive domain knowledge was needed. In a short conversation the experiment was explained to the participants beforehand. The main focus of this introduction was the structure of the annotations and the underlying ontology, further the domain of Tiny Planets was briefly introduced. There was no time limit set for the completion of the experiment.

In the first run, the query “Bing telescope” was given to the proband together with 8 images. In the second run the query “Alien smiling” was given to the proband and 7 images. Finally in the third run the query “Bing umbrella” was given to the proband and 6 images.

The human results are compared on the one hand with the results of the initial ranking approach and on the other hand with the refined ranking approach. Spearman’s rank correlation coefficient can be used to determine the correlation between the rankings of the algorithms with the one’s done by the propands. Therefore for each proband’s result a correlation between the refined approach and the initial approach is computed. The correlation is a value between the interval -1 and 1. Where 1 is a complete correlation and -1 means one ranking is the reverse of the other. With this measure the effectiveness of the two ranking algorithms can be compared. If the correlation for the refined approach is higher than the one for the initial approach, the refined one is better. If the correlation for the refined approach is lower than the one for the initial approach, the initial approach is better. If the correlation for the two approaches is the same, they are equal.

Results

In the following table the results of the experiment are presented. On the left hand side of each table the resources are listed, each resource is represented as a capital letter. In the following columns the ranking of the two algorithms and of the test persons is shown. A 1 represents the highest rank. The last two rows in each table show the correlations of the initial algorithm with all rankings of the test persons and the correlations of the refined approach with all the test person’s rankings. It can be easily observed that in all cases the correlation of the refined approach is higher than the correlation of the initial approach. According to the results of the experiment the refinements have improved the algorithm.

Query 1 : "Bing Telescope"							
resources	refined approach	initial approach	person 1	person 2	person 3	person 4	person 5
A	1	1	2	2	2	2	3
B	8	7	8	8	7	6	8
C	3	8	3	3	3	4	2
D	5	2	4	7	5	3	6
E	4	3	5	5	6	8	7
F	6	5	6	4	4	5	4
G	7	6	7	6	8	7	5
H	2	4	1	1	1	1	1
correlation refined approach			0.952	0.857	0.857	0.667	0.714
correlation initial approach			0.452	0.214	0.310	0.357	0.000

Query 2 : "Alien Smiling"							
resources	refined approach	initial approach	person 1	person 2	person 3	person 4	person 5
A	5	4	4	5	4	4	5
B	4	1	5	4	5	5	4
C	1	2	1	1	1	1	1
D	7	7	6	6	6	7	7
E	3	3	3	3	3	3	3
F	6	6	7	7	7	6	6
G	2	5	2	2	2	2	2
correlation refined approach			0.929	0.964	0.929	0.964	1.000
correlation initial approach			0.500	0.607	0.500	0.536	0.643

Query 3 : "Bing Umbrella"							
resources	refined approach	initial approach	person 1	person 2	person 3	person 4	person 5
A	2	2	3	2	1	2	2
B	3	3	2	3	3	3	3
C	1	6	1	1	2	1	1
D	5	5	4	4	4	6	5
E	4	1	6	5	6	4	6
F	6	4	5	6	5	5	4
correlation refined approach			0.771	0.943	0.771	0.943	0.771
correlation initial approach			-0.543	-0.314	-0.257	-0.029	-0.429

Table 1: Experimental Results

Discussion

Firstly it has to be mentioned that a ranking of semantically annotated resources can only be as good as the annotations provided. If the annotations of a multimedia resource are not representing the content correctly, then the ranking cannot be effective.

The experimental evaluation of the refinements presented in the previous section cannot proof the increase of effectiveness regarding to the initial approach as a panel of 5 humans is typically not a critical mass. However the experimental results show a clear tendency that the results are ranked closer to a user's expectations than in the initial approach.

As an information retrieval system can only retrieve the relative best answers to a query a ranking algorithm can never be perfect. Currently there is already one aspect of the underlying resource annotations which is not taken into consideration by the algorithm. As briefly mentioned in Section 3 relationships between resources can be created. Therefore if a resource is popular, which means it has many in- and outgoing relationships it could be more relevant in some situations. As we have not yet found any evidence for this assumption it has not been considered.

The refined ranking approach is not a general purpose approach. The algorithm is quite specific to the underlying ontology. The algorithm is quite dependent triples of the form: *<concept x> <related to> <concept y>*. If the constraining predicate "related to" could be made more flexible, a more general purpose practicability can be achieved.

5.5 Integration

The facilities of the IMAS to annotate and search media items are available for other applications. Figure 13 illustrates the integration of applications and services. The AspectBrowser, part of the content-based search, delegates requests to annotate media items to the IMAS by accessing it with a URI parameter. To search media items, the content-based search application integrates the search facilities of the content-based service and the semantic service on its own, because it uses a third data source, the WWW. Further details and the results of this integration are documented in Deliverable D5.5.4

The FBM-UPF's Program Editor uses the IMAS to annotate and search media items. It is possible to create annotations of media items by selcting the "Annotation" entry inside in a context-menu of the Program Editor. This action opens a browser with IMAS to annotate the previously selected media item. To search and find suitable media item, users can use the search facilites of IMAS within the Program Editor.

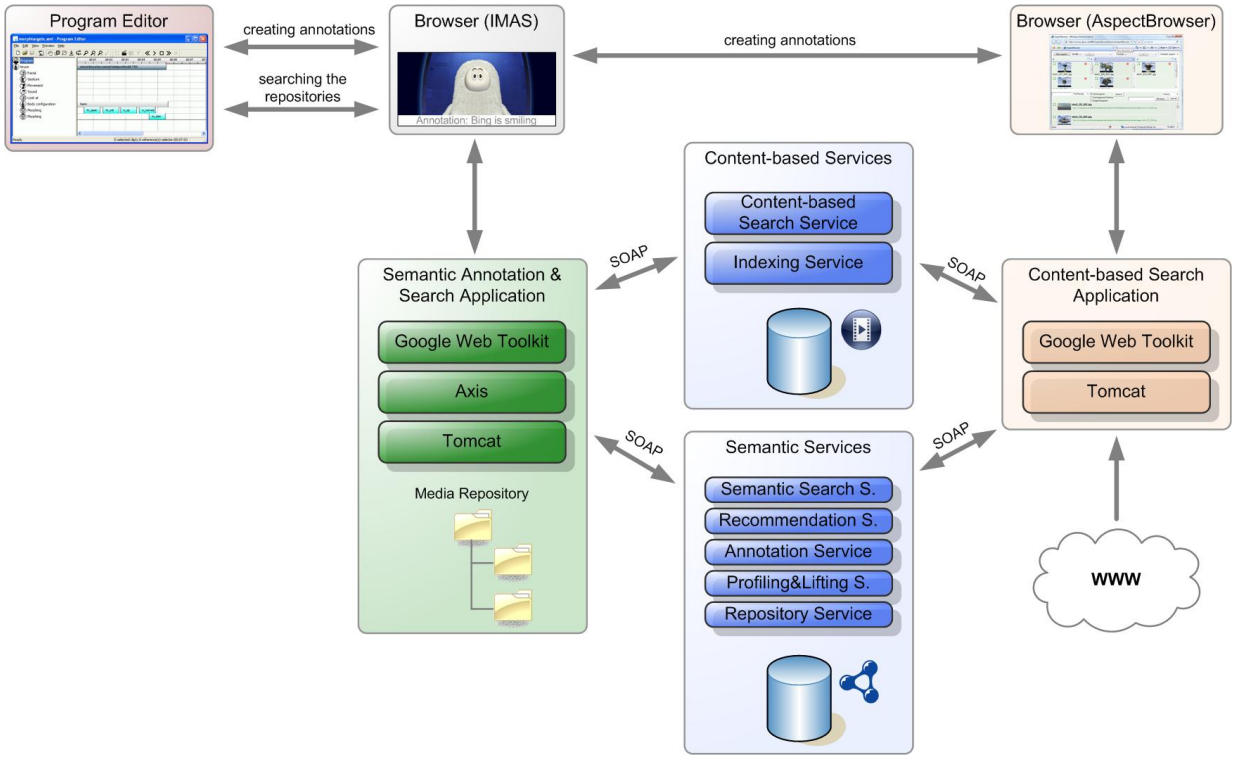

Figure 13: IMAS integration overview

6 Evaluation

6.1 Introduction

This chapter describes the evaluation of annotation functionality of the Semantic Annotation Tool which took place in April 2009.

Nine test users had to annotate a set of ten images using the SALERO Semantic Annotation Tool. As a comparison the same persons had to complete the same task with Google's Picasa [Picasa] software.

The test users agreed that the automatic suggestion / completion of tags of the SALERO Semantic Annotation Tool was very helpful in the annotation process. The concept of statement based annotation was not clear to all users, the length of statements ranged from short to very long. However, some users missed the possibilities to enter structured annotations when using Google Picasa.

When asked after using both tools, which one they would use for annotation a larger collection (1000 images) annotation tool most users would choose the SALERO Semantic Annotation Tool.

As a result of the evaluation process a set of recommendations was defined, how the user interface will be further developed to streamline the annotation process.

This chapter further discussed the feedback gathered from the attendants of the fourth user group meeting held in Annecy

6.2 Goal of the Evaluation

Our aim was (i) to find major obstacles in the annotation user interface as well as in the annotation process and (ii) to compare the semantic statement-based approach with other existing approaches.

The search functionality of the tool was not object of this evaluation.

6.3 Methodology

We recruited 9 participants of our institute who were not involved in the project or in the realm of semantic annotation. The subjects ranged in age from 25 to 40 and all are software developers. We created two groups. The first group of five participants compared the annotation process of the IMAS application with the desktop application of Google Picasa [Picasa]. In the second group we did chronometries of creating annotations with (i) IMAS, (ii) a free text tagging approach, similar to flickr[17] and (iii) creating fully semantic annotations with PhotoStuff [Halaschek-Wiener et al., 2005]. For later analysis we did screen captures of the user actions and conspicuous behaviour was noted by the observer. Furthermore, the usability test consisted of a short user questionnaire where the participants had to answer following questions:

1. What are the positive and negative impressions of each tool?
2. What was the most difficult task?
3. Which tool would you use, if you have to annotate 1000 images?

Before the test began, each subject got an introduction in the usability test and into the different annotation applications. Then the users had to annotate the test data with each application. The test data consisted of ten images from the tiny planets universe together with descriptive text for each image. For example, the descriptive text for Figure 7 is: "Bing has a bag on his shoulder and a book in his hands. Bong and alien are smiling."

[17] Flickr: http://www.flickr.com/

6.4 Test Environment

6.4.1 *Location, Time and Involved Personnel*

- Location: Office of Philip Hofmair, JOANNEUM RESEARCH, Steyrergasse 17, Graz, Austria.
- Date: 29 April, 2009
- Time: 10 am – 5 pm
- Evaluator: Hofmair Philip (JRS)
- Assistants: Gert Kienast (JRS) and Wolfgang Weiss (JRS)

6.4.2 *System Environment*

- PC: Dell Optiplex 755, 2GB RAM, Core2Duo E8400 @ 3GHz
- Dell TFT Display @ 1600x1200 resolution
- MS Windows XP Pro Version 2002, Service Pack 3
- Firefox 3.0.10
- SALERO Semantic Annotation Tool (version of 2009/04)
- Google Picasa 3

6.4.3 *Test Data*

- 10 images with description of content (see Annex I)

6.4.4 *Overview of the Test Persons*

- 9 male persons, 25 to 40 years old
- Researchers / Developers at JRS
- Education: post graduate
- Usually used software: MS Windows, Office, Outlook; IDE's for Java and C++
- Internet / Computer usage: 25+ hours per week

6.5 Evaluation Results

The success rate to complete the annotation tasks with IMAS is 93%. The reasons why not all participants were able to successfully complete all tasks are (i) user interface and design problems and (ii) that the semantic statement-based approach was not clear for every participant and thus produced wrong or incomplete annotations. A successful example, done by a participant with IMAS, Includes following statements for the image Figure 7:

- Bing is related to: Shoulder bag, Shoulder
- Bing is related to: Book, Hand
- Bong is related to: smiling
- Alien is related to: smiling

and the following tags created with Google Picasa: "alien, bing, bong, book, hand, shoulder bag on his shoulder, smiling". These examples describe the content according the descriptive text of the image well and fulfil our quality requirements for a successful annotation. Figure 14 shows the time measurements for completing the tasks where the statement-based approach is compared to a free text tagging approach and to creating fully semantic annotations with PhotoStuff. The original data for this figure can be found in section 13. Creating annotations with PhotoStuff was the most time consuming approach (median 60s; mean 69.9s for creating the annotations of a single media resource). The subjects

complained that they have to do long-winded recurrent tasks, such as selecting concepts and manually creating instances. The fastest approach was to use the simple free text tagging approach (median 19s; mean 18.6s), although the subjects claimed to use a system with auto completion feature. Task completion time using IMAS with the statement-based approach ranks between the two other systems with a median of 30s and a mean of 36.3s.

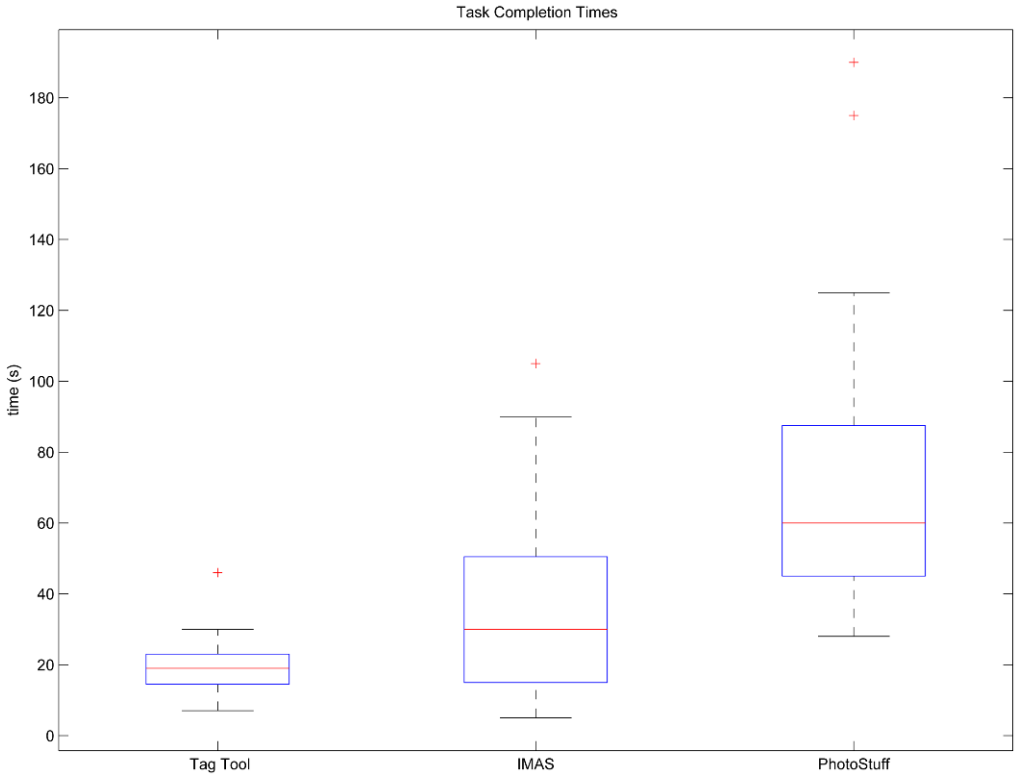

Figure 14: Comparison of task completion times.

We observed that most subjects firstly used the concept tables (see also Figure 8 and section 5.2.1) and after annotating approximately three media resources the subjects tended to use only the command line interface with auto completion. The questionnaire revealed following facts: The users liked the auto completion feature which demonstrates following user statement:

> *"I highly appreciate that the system suggests concepts to me I can use."*

In the subjects opinion this feature helps efficiently to create specific suitable annotations. Furthermore, this was a crucial reason why for 8 out of 9 participants IMAS is the first choice for annotating 1000 media resources. One participant prefers to use a simple free text tagging approach. On the other hand, the participants also revealed new suggestions to improve the annotation process:

> *"A copy and paste functionality of concepts and statements would be fine."*

> *"Existing annotations should be editable."*

The users also complained that the text box does not work as expected in some situations, e.g. when pressing the space bar to select a concept.

6.6 Detailed Analysis and Notes

6.6.1 System 1 (SALERO Semantic Annotation Tool)

Analysis

- Firefox drag & drop problems. Concepts have to be clicked twice before they can be dragged.
- The concept tables are mostly used as an overview of available concepts and not for annotation itself. Most users prefer to just type in text field.

- The concept of statements is not clear to all users. The participants are not sure to do the right thing.
- Editing of existing statements was a wish of some participants. At the moment the statement has to be deleted and created new.
- Space Key was used for confirming a concept when typing in the text field → Problem when more than one matching concept was found.
- Use the auto completion mechanism also in the concept table filter functionality. Image #3 shows a couch/sofa and the descriptive text says “couch”. Typing in the name couch in the table filter ends up with no result. Typing in couch in the text field ends up with the concept sofa in the suggestion popup.
- „Selected Media Items” were misunderstood.
- Browser Back – Functionality with the back space key lead to a loss of the entered annotation text.
- Multi-select of images is misleading (in the selected media items area)
- One participant tried to edit a statement via dragging it to the text field and was not successful because this functionality is not implemented.

User Comments

- “Do not show dialog after saving ‘Data saved successfully.’ Show a message when saving was not successful instead.”
- “A Copy & Paste functionality of concepts and statements would be fine!“
- Recently used annotations are missing
- Annotation and file browser should be located in one Tab.
- “The system shows me concepts I can use for annotation – very nice!”
- “I would prefer the Semantic Annotation Tool to annotate a dataset of 1000 images than Picasa!”
- “Automatic detection of concepts in a image would be nice.”
- If a concept is recognised space and or enter key should confirm the concept.
- Existing annotations should be editable.

6.6.2 System 2 (Google Picasa)

Analysis

- Annotation of images via thumbnails works good. Participant selects more than one image and tags them.
- Thumbnail view of images is much better than the tree view.
- Tagging is only a flat list of concepts – why should I do it?
- I have to think of concept names myself. I don’t like it.
- Some navigation problems occurred when participants double clicked an image. They did not find the way back to the previous view.

User Comments

- + Tag proposal based on existing tags.
- + The tags are editable
- + overall handling of images is easy

6.7 Recommendations

In a follow-up discussion (bi-weekly "SALERO breakfast" at JRS on 4 May 2009) with the entire JRS SALERO team, the following actions were prioritized:

- Redesign User Interface:
  - Have the file browser and annotation tab on one screen.
  - This will also solve ambiguities when selecting multiple images.
- Allow SPACE-key to select highlighted concept during auto-completion and add menu entry "New concept" to suggestion box
- Allow editing of existing annotations
- Disable Back-function in Browser when hitting the Backspace-key
- Remove message "Data saved successfully". Only show message when error occurs.
- Show information about statement concepts on demand
- Extend the filter functionality in the concept tables
- Thumbnail view of all selected media files in the preview panel of the annotation tab

6.8 Feedback from the SALERO User Group

In the feedback form filled out by the attendants of the fourth user group meeting (at the International Animation Film Festival in Annecy, France, June 2009) the following feedback was gathered. Questions 7, 8, 9 and 10 were targeted at their perception of Semantic Annotation and Search.

Ten of the twelve attendants who returned the questionnaire answered the questions on the Semantic Annotation and Search Tool. Two have not answered the question because the have not seen the demo due to arriving late or leaving early.

The attendants had the choice between four answers for every question and had the possibility to add textual comments on every question. The following section lists the answers and comments given by the user group.

The full questionnaire results can be found in Annex III of this document.

6.8.1 Answers & comments concerning Semantic Annotation and Search

Q7: Was the approach of annotating media resources easy and comprehensible?		
Very easy	1	(10%)
Rather easy	5	(50%)
Rather difficult	4	(40%)
Very difficult	0	(0%)
Total	*10*	*(100%)*

Comments received on Q7:

- Clear and well structured. Will help to save time in the process.
- While the interface convention adopted was straightforward enough, a more intuitive approach using drag and drop methodology, for example, would be less tedious and would simplify the tagging process. Also, tools to customize tags for project-specific relevance would save time in the search process.
- This was an application that would be very useful for larger projects with many artists working on a series. I found the presenter very open to ideas for making it more intuitive for artists.

- Text input is fine for most, but drag and drop would be very useful on the menus
- Creating the list of materials is being the most precise and time consuming work
- The system crashed on demonstration, and seemed difficult to explain.

Q8-1: Could you please tell us how important are the following factors for a semantic annotation tool: **Seamless integration into other (authoring) applications, e.g. 3ds Max, Photoshop, Cinema 4D…**		
Very important	2	(20%)
Rather important	6	(60%)
Marginally important	2	(20%)
Not important	0	(0%)
Total	*10*	*(100%)*

Q8-2: Could you please tell us how important are the following factors for a semantic annotation tool: **To find again media resources created in previous projects**		
Very important	6	(60%)
Rather important	3	(30%)
Marginally important	0	(0%)
Not important	1	(10%)
Total	*10*	*(100%)*

Q8-3: Could you please tell us how important are the following factors for a semantic annotation tool: **Possibility to select one of several meanings in case of homonyms (e.g. bank (financial institution) vs. bank(seating))**		
Very important	2	(20%)
Rather important	5	(50%)
Marginally important	0	(0%)
Not important	3	(30%)
Total	*10*	*(100%)*

Comments received on Q8-1, -2 and -3:

- The main thing is that content can be easily retrieved by someone other than the person who carried out the initial tagging. Support for homonyms is far less relevant than support for misspelling.
- It might be more useful to have thumbnails and a glossary to help the artist find assets more easily.
- It would be useful in making filming and production notes especially. We could organise thoughts on camera settings for stereo and the like, as long as it was fairly seamless to use.
- Well organized and detailed data bank is important factor in all projects where previously used data could be reused, by having an easy access to all information needed.
- Probably there would be time when musicians would look for specific material they recorded some time ago… but nonetheless they would not start filing all the material for that purpose.

Q9: Could you please tell us how relevant to your business the tools demonstrated is?		
Very relevant	2	(20%)
Relevant	2	(20%)
Marginally relevant	4	(40%)
Not relevant	2	(20%)
Total	*10*	*(100%)*

Comments received on Q9:

- In principle, this tool has great potential for managing media intelligently. The tool still requires refinement before it would be suitable for use in a commercial environment. The developer is clearly very responsive to input from potential users.
- It would be useful in the case of series work, after the initial pilot were sold.
- It is likely one of the things that would need to be used for a few months before its true value is recognized, but the ability to amass stereo filming knowledge and stream that out in near real-time to a database or website would be excellent.
- This seemed a lot of work for not much return

Q10: Please, score the cost impact in your business if using this tool.		
Costs would decrease	1	(14%)
Costs would be the same	2	(29%)
Costs would increase	2	(29%)
Costs would increase, although it would be compensated for a better quality of service	2	(29%)
Total	*9*	*(100%)*

Comments received on Q10:

- There could be an increase in efficiency in regards to analyzing filming techniques after and before jobs.
- Would the material be listed automatically – like, all Shakespeare books scanned in and then it would be easy to look up a specific quote or thought… But to make specific filing a day to day job in order to find a part in years to come – this sounds more like a job for libraries than everyday other business. Too work consuming therefore costly for my use.
- Once ease of use and project-specific relevance features improved, this tool would be an excellent addition to multimedia project development.
- If some of the changes we suggested were implemented (and tags could be created for specific individual needs) it would be very useful in organizing re-use of assets after the original pilot scenes were completed, and then added to as the project continued.

6.9 SWOT Analysis

Strengths: • Semantic annotations • Controlled vocabulary • Simple and fast • Integration of different search engines (semantic search & content-based search) • Competitive lead in annotation and search of media resources	**Weaknesses:** • Integration with UG's tools is currently implemented as a distributed System (e.g. content-based search system is installed on a separate server). This needs further attention.
Opportunities: • Real life installation at a industry partner	**Threats:** • Benefits of the tool will not pay off immediately, but only in later projects. Hence production companies may be reluctant to start using the tool, because annotation requires additional work (maybe under heavy work load and deadlines closing in) in the first place.

7 Conclusion

We have described a semantic statement-based approach for fast and easy annotation of media resources in the realm of media productions. This concept is implemented in the Intelligent Media Annotation & Search[18] application which also allows creating relationships of media resources.

Creating annotations using statements with semantic elements is a compromise in complexity and expressiveness between loose and full semantically descriptions. We have developed two ontologies to store annotations and relationships.

The annotation and search functionality has been integrated with partners' tools, namely the context-based search (UG) and the Program Editor (FBM-UPF).

The system integrates semantic and content-based search to provide a fall-back and an alternative retrieval system.

The initial usability test has shown that the approach of semantic statement-based annotation is not as fast as free text tagging, but much faster than creating full semantic annotations with PhotoStuff.

[18] IMAS Online Prototype: http://salero.joanneum.at/imas/

8 References

[Andrews, 2009] K. Andrews, Human-Computer Interaction, Lecture Notes, Version of 28 Apr 2009, Chapter 8+9, University of Technology, Graz, Austria http://courses.iicm.tugraz.at/hci/hci.pdf (last access 7 August 2009)

[Bürger et al. 2006] T. Bürger, D. Denissenko, A. Wahler, H. Zeiner, G. Kienast, R. Villa, C. Cullen, T. Paju, and J. Blat "Ontology Language for Multimedia Objects" SALERO Deliverable D3.1.2, 2006.

[Bürger et al. 2007] T. Bürger, E. Simperl, G. Kienast, P. Hofmair, and H. Zeiner "Ontologies for Multimedia Objects and Workbench (Initial Version)" SALERO Deliverable 3.1.3.

[Bürger et al., 2008] T. Bürger, P. Hofmair, and G. Kienast „The SALERO Virtual Character Ontology“ In: Proceedings of the First Workshop on Semantic 3D Media, co-located with SAMT 2008, Dec, 3.-5., 2008

[Bürger et al., 2009] T. Bürger, G. Kienast, W. Weiss, R. Villa, P. Swamy, D. Hannah. „Embedding of Semantic Metadata in Content Formats” SALERO Deliverable D3.1.6, 2009, Available online: http://www.salero.eu/media/pdf/del/SALERO-D3.1.6-EmbeddingOfSemanticMetadataInContentFormats.pdf

[Bürger and Ammendola, 2008] Bürger, T., Ammendola, C.: A user centered annotation methodology for multimedia content. In: Poster Proceedings of ESWC 2008. (2008)

[Fernández-López et al, 1999] Fernández-López, M., Gómez-Pérez, A., Pazos, A., & Pazos, J. (1999). Building a chemical ontology using methontology and the ontology design environment. IEEE Intelligent Systems & Their Applications, 4(1), 37-46.

[FRBR, 2009] IFLA Study Group: Functional requirements for bibliographic records, final report. http://www.ifla.org/files/cataloguing/frbr/frbr_2008.pdf (February 2009).

[Hoi and Lyu, 2008] Hoi S. C. H., Lyu M. R.: A multimodal and multilevel ranking scheme for Large-Scale Video Retrieval. In: IEEE transactions on Multimedia. Vol. 10, No. 4, 607 -619, (2008).

[Grüninger et al, 1995] Grüninger, M. & Fox, M.S. (1995). Methodology for the design and evaluation of ontologies. In D. Skuce (Ed.), IJCAI95 Workshop on Basic Ontological Issues in Knowledge Sharing (pp. 6.1–6.10).

[Halaschek-Wiener et al., 2005] Halaschek-Wiener, C., Golbeck, J., Schain, A., Grove, M., Parsia, B., Hendler, J.A.: PhotoStuff - An Image Annotation Tool for the Semantic Web. In: Poster Proceedings of the 4th International Semantic Web Conference. (2005)

[Kerrigan et al. 2007] Kerrigan, M., Mocan, A., Tanler, M., Fensel, D.: The web service modeling toolkit - an integrated development environment for semantic web services (system description). In: Proceedings of the 4th European Semantic Web Conference (ESWC), June 2007, Innsbruck, Austria. (2007)

[Kienast et al. 2008] G. Kienast, P. Hofmair, T. Bürger, C. Ammendola, M. Nieminen, K. Åberg, and M. Romeo "Ontologies for Multimedia Objects and Workbench – Update" SALERO Deliverable D3.1.5.

[Kludas et al., 2008] Kludas, J., Bruno, E., Marchand-Maillet S.: Can Feature Information Interaction help for Information Fusion in Multimedia Problems?. In: Proc. Of the first International Workshop on Metadata Mining for Image Understanding (MMIU), Funchal, Madeira, (2008).

[Mc Donald and Smeaton, 2005] Mc Donald, K., Smeaton A. F.: A comparison of score, Rank and Probability-Based Fusion Methods for Video Shot Retrieval. In: In Proc of Intl. Conf on Image and Video Retrieval. 61-70, (2005).

[O'Connor et al., 2005] O'Connor, N., Cooke, E., le Borgne, H., Blighe, M., Adamek, T.: The ace-toolbox: low-level audiovisual feature extraction for retrieval and classification. In: 2nd IEE European Workshop on the Integration of Knowledge, Semantic and Digital Media Technologies. (2005) 55-60

[Ounis et al., 2006] Ounis, I., Amati, G., Plachouras, V., He, B., Macdonald, C., Lioma, C.: Terrier: A High Performance and Scalable Information Retrieval Platform. In: Proc. of ACM SIGIR'06 Workshop on Open Source Information Retrieval (OSIR). (2006)

[Picasa] Google Picasa, http://picasa.google.com

[Raskin, 2008] Raskin, A.: The Linguistic Command Line. In: ACM Interactions. (January and February 2008) 19-22

[Riva, 2007] Riva, P.: Introducing the functional requirements for bibliographic records and related ifla developments. ASIS&T The American Society for Information Science and Technology (August/September 2007).

[Noy et al, 2001] N.F. Noy and D.L. McGuinness. “Ontology Development 101: A Guide to Creating Your First Ontology". Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001.

[Staab et al., 2001] Staab, S., Schnurr, H.P., Studer, R., & Sure, Y. (2001). Knowledge processes and ontologies. IEEE Intelligent Systems, 16(1), 26-34.

[Qi and Han, 2005] Qi, X., Han, Y.: A novel fusion approach to content based image retrieval. In: Pattern Recognition. Vol.38, 2449 -2465, (2005).

[Uschold et al., 1995] Uschold, M., & King, M. (1995). Towards a methodology for building ontologies. In D. Skuce (Ed.), IJCAI’95 Workshop on Basic Ontological Issues in Knowledge Sharing (pp. 6.1-6.10), Montreal, Canada.

[Westerveld and Vries, 2004] Westerveld, T., de Vries A.P.: Multimedia Retrieval using Multiple Examples. In: LNCS Image and Video Retrieval. Vol. 3115, (2004) 344-352.

9 Glossary

Partner Acronyms

AM	Activa Multimedia, ES
BLITZ	Blitz Games, UK
DFT	Digital Film Technology, DE
DIT	Dublin Institute of Technology, IE
DLLNI	DTS Licensing Ltd. Northern Ireland, UK
FBM-UPF	Fundació Universitat Pompeu Fabra, ES
FUNITEC	Universitat Ramon Llull, ES
JRS	JOANNEUM RESEARCH Forschungsgesellschaft mbH, AT
LFUI	Leopold-Franzens Universität Innsbruck, AT
PGP	Pepper's Ghost Productions Ltd., UK
TAIK	Taideteollinen Korkeakoulu, FI
UG	University of Glasgow, UK
UPF	Universitat Pompeu Fabra, ES

10 Annex I: Evaluation Test Data

Test data for the user interface evaluation consisted of 10 images from the TinyPlanets universe together with some descriptive text. Users were asked to annotate the given text using the SALERO Semantic Annotation Tool and Google Picasa.

Image	Ground truth annotation
	Bing has a bag on his shoulder and looks through a telescope while Bong is smiling.
	Bing has a bag on his shoulder and a book in his hands. Bong and the alien are smiling.
	Bing and Bong are sitting on a couch and flying through the space. Bong is controlling the steering wheel and both have fastened the seatbelt.
	This scene illustrates Bing, Bong and an alien. Bing has a bag on his shoulder and is waving his hand while Bong is smiling. There is also a chair and a table with an ice cream on it.

	Bing is ringing the doorbell of the alien's house. The alien stands in front of Bing.
	Bong is smiling in this scene.
	Bing and Bong are sleeping and dreaming in a Bed.
	Bong is smiling and has a pencil in his mouth.
	The alien looks surprised and has a boot on his leg.

	The alien, in front of a broken egg shell, looks surprised.

11 Annex II: Report Data Sheet (German)

SALERO Usability Test

Bericht

PersonNr. _____

Datum _____

Tool 1 (Semantic Annotation Tool)	Tool 2 (Picasa3)
B+B.JPG: Aufgabe erfüllt: JA / NEIN Anmerkung:	**B+B.JPG:** Aufgabe erfüllt: JA / NEIN Anmerkung:
B+B_Pose_Extra001_01_Export.jpg: Aufgabe erfüllt: JA / NEIN Anmerkung:	**B+B_Pose_Extra001_01_Export.jpg:** Aufgabe erfüllt: JA / NEIN Anmerkung:
B+B_Pose_TP059_04.jpg: erfüllt: JA / NEIN Anmerkung:	**B+B_Pose_TP059_04.jpg:** erfüllt: JA / NEIN Anmerkung:
Bing&Bong01.jpg: Aufgabe erfüllt: JA / NEIN Anmerkung:	**Bing&Bong01.jpg:** Aufgabe erfüllt: JA / NEIN Anmerkung:
Bing&Bong03.jpg: Aufgabe erfüllt: JA / NEIN Anmerkung:	**Bing&Bong03.jpg:** Aufgabe erfüllt: JA / NEIN Anmerkung:
Bing&Bong02.jpg : Aufgabe erfüllt: JA / NEIN Anmerkung:	**Bing&Bong02.jpg :** Aufgabe erfüllt: JA / NEIN Anmerkung:
Bing&Bong06.jpg: Aufgabe erfüllt: JA / NEIN Anmerkung:	**Bing&Bong06.jpg:** Aufgabe erfüllt: JA / NEIN Anmerkung:
bong.jpg: Aufgabe erfüllt: JA / NEIN Anmerkung:	**bong.jpg:** Aufgabe erfüllt: JA / NEIN Anmerkung:
F_Pose_Extra001_01.jpg: erfüllt: JA / NEIN Anmerkung:	**F_Pose_Extra001_01.jpg:** erfüllt: JA / NEIN Anmerkung:

F_Pose_Extra001_04.jpg: erfüllt: JA / NEIN Anmerkung:	**F_Pose_Extra001_04.jpg:** erfüllt: JA / NEIN Anmerkung:

Positiv aufgefallen:

Tool 1 (Semantic Annotation Tool)	Tool 2 (Picasa3)

Negativ aufgefallen:

Tool 1 (Semantic Annotation Tool)	Tool 2 (Picasa3)

Welches Tool sagt dir am ehesten zu, bzw. welches würdest du verwenden wenn du 1000 Dateien annotieren müsstest?

12 Annex III: User Group Questionnaire Results

This annex lists the result from questionnaire distributed to the User Group after the meeting in Annecy (this data is used in section 6.8.1). From each questionnaire the questions related to semantic tools (number 7 to 10) have been extracted

12.1 Person 1

7) Was the approach of annotating media resources easy and comprehensible?
[ ] very easy [ ] rather easy [X] rather difficult [ ] very difficult

Comments:
……

8) Could you please tell us how important are the following factors for a semantic annotation tool?

Seamless integration into other (authoring) applications, e.g. 3ds Max, Photoshop, Cinema 4D…
[ ] very important [X] important [ ] marginally important [ ] not important

To find again media resources created in previous projects:
[X] very important [ ] important [ ] marginally important [ ] not important

Possibility to select one of several meanings in case of homonyms (e.g. bank (financial institution) vs. bank(seating))
[ ] very important [X] important [ ] marginally important [ ] not important

Comments:
……

9) Could you please tell us how relevant to your business the tools demonstrated is?

[ ] very relevant [ ] relevant [ ] marginally relevant [X] not relevant

Comments:
……

10) Please, score the cost impact in your business if using this tool:

[ ] Costs would be decremented
[ ] Costs would be the same
[ ] Costs would be increased
[ ] Costs would be increased, although it would be compensated for a better quality of service

Comments:
……

12.2 Person 2

7) Was the approach of annotating media resources easy and comprehensible?
[ ] very easy [*] rather easy [ ] rather difficult [ ] very difficult

Comments:
While the interface convention adopted was straightforward enough, a more intuitive approach using drag and drop methodology, for example, would be less tedious and would simplify the tagging process. Also, tools to customize tags for project-specific relevance would save time in the search process.
……

8) Could you please tell us how important are the following factors for a semantic annotation tool?

Seamless integration into other (authoring) applications, e.g. 3ds Max, Photoshop, Cinema 4D…
[ ] very important [*] important [ ] marginally important [ ] not important

To find again media resources created in previous projects:
[*] very important [ ] important [ ] marginally important [ ] not important

Possibility to select one of several meanings in case of homonyms (e.g. bank (financial institution) vs. bank(seating))
[ ] very important [ ] important [ ] marginally important [*] not important

Comments:
The main thing is that content can be easily retrieved by someone other than the person who carried out the initial tagging, Support for homonyms is far less relevant than support for misspelling.
……

9) Could you please tell us how relevant to your business the tools demonstrated is?

[*] very relevant [ ] relevant [ ] marginally relevant [ ] not relevant

Comments:
In principle, this tool has great potential for managing media intelligently. The tool still requires refinement before it would be suitable for use in a commercial environment. The developer is clearly very responsive to input from potential users.
……

10) Please, score the cost impact in your business if using this tool:

[ ] Costs would be decremented
[ ] Costs would be the same
[ ] Costs would be increased
[*] Costs would be increased, although it would be compensated for a better quality of service

Comments:
Once ease of use and project-specific relevance features improved, this tool would be an excellent addition to multimedia project development.
……

12.3 Person 3

7) Was the approach of annotating media resources easy and comprehensible?
[ ] very easy [ ] rather easy [X] rather difficult [ ] very difficult

Comments:

The system crashed on demonstration, and seemed difficult to explain.
……

8) Could you please tell us how important are the following factors for a semantic annotation tool?

Seamless integration into other (authoring) applications, e.g. 3ds Max, Photoshop, Cinema 4D…
[ ] very important [X] important [ ] marginally important [ ] not important

To find again media resources created in previous projects:
[X] very important [ ] important [ ] marginally important [ ] not important

Possibility to select one of several meanings in case of homonyms (e.g. bank (financial institution) vs. bank(seating))
[X] very important [ ] important [ ] marginally important [ ] not important

Comments:
……

9) Could you please tell us how relevant to your business the tools demonstrated is?

[ ] very relevant [X] relevant [ ] marginally relevant [ ] not relevant

Comments:
……

10) Please, score the cost impact in your business if using this tool:

[X] Costs would be decremented
[ ] Costs would be the same
[ ] Costs would be increased
[ ] Costs would be increased, although it would be compensated for a better quality of service

Comments:
……

12.4 Person 4

7) Was the approach of annotating media resources easy and comprehensible?
[ ] very easy [X] rather easy [ ] rather difficult [ ] very difficult

Comments:
…This was a application that would be very useful for larger projects with many artists working on a series. I found the presenter very open to ideas for making it more intuitive for artists.
……

8) Could you please tell us how important are the following factors for a semantic annotation tool?

Seamless integration into other (authoring) applications, e.g. 3ds Max, Photoshop, Cinema 4D…
[ ] very important [X] important [ ] marginally important [ ] not important

To find again media resources created in previous projects:
[X] very important [ ] important [ ] marginally important [ ] not important

Possibility to select one of several meanings in case of homonyms (e.g. bank (financial institution) vs. bank(seating))
[ ] very important [ ] important [ ] marginally important [X] not important

Comments:
…It might be more useful to have thumbnails and a glossary to help the artist find assets more easily.
……

9) Could you please tell us how relevant to your business the tools demonstrated is?

[X] very relevant [ ] relevant [ ] marginally relevant [ ] not relevant

Comments:
…It would be useful in the case of series work, after the initial pilot were sold.
……

10) Please, score the cost impact in your business if using this tool:

[ ] Costs would be decremented
[ ] Costs would be the same
[ ] Costs would be increased
[X] Costs would be increased, although it would be compensated for a better quality of service

Comments:
…If some of the changes we suggested were implemented (and tags could be created for specific individual needs) it would be very useful in organizing re-use of assets after the original pilot scenes were completed, and then added to as the project continued.
……

12.5 Person 5

7) Was the approach of annotating media resources easy and comprehensible?
[ ] very easy [x] rather easy [ ] rather difficult [ ] very difficult

Comments:
……

8) Could you please tell us how important are the following factors for a semantic annotation tool?

Seamless integration into other (authoring) applications, e.g. 3ds Max, Photoshop, Cinema 4D…
[ ] very important [x] important [ ] marginally important [ ] not important

To find again media resources created in previous projects:
[x] very important [ ] important [ ] marginally important [ ] not important

Possibility to select one of several meanings in case of homonyms (e.g. bank (financial institution) vs. bank(seating))
[ ] very important [x] important [ ] marginally important [ ] not important

Comments:
……

9) Could you please tell us how relevant to your business the tools demonstrated is?

[ ] very relevant [x] relevant [ ] marginally relevant [ ] not relevant

Comments:
……

10) Please, score the cost impact in your business if using this tool:

[x] Costs would be decremented
[ ] Costs would be the same
[ ] Costs would be increased
[ ] Costs would be increased, although it would be compensated for a better quality of service

Comments:
……

12.6 Person 6

Did not fill out this part of the questionnaire.

12.7 Person 7

7) Was the approach of annotating media resources easy and comprehensible?
[ ] very easy [x] rather easy [ ] rather difficult [ ] very difficult

Comments:
Text input is fine for most, but drag and drop would be very useful on the menus……

8) Could you please tell us how important are the following factors for a semantic annotation tool?

Seamless integration into other (authoring) applications, e.g. 3ds Max, Photoshop, Cinema 4D…
[x] very important [ ] important [ ] marginally important [ ] not important

To find again media resources created in previous projects:
[x] very important [ ] important [ ] marginally important [ ] not important

Possibility to select one of several meanings in case of homonyms (e.g. bank (financial institution) vs. bank(seating))
[x] very important [ ] important [ ] marginally important [ ] not important

Comments:
It would be useful in making filming and production notes especially. We could organise thoughts on camera settings for stereo and the like, as long as it was fairly seamless to use……

9) Could you please tell us how relevant to your business the tools demonstrated is?

[ ] very relevant [ ] relevant [x] marginally relevant [ ] not relevant

Comments:
It is likely one of the things that would need to be used for a few months before its true value is recognized, but the ability to amass stereo filming knowledge and stream that out in near realtime to a database or website would be excellent.…….

10) Please, score the cost impact in your business if using this tool:

[ ] Costs would be decremented
[x] Costs would be the same
[ ] Costs would be increased
[ ] Costs would be increased, although it would be compensated for a better quality of service

Comments:
There could be an increase in efficiency in regards to analyzing filming techniques after and before jobs.…….

12.8 Person 8

7) Was the approach of annotating media resources easy and comprehensible?

[ ] very easy [ ] rather easy [ ] rather difficult [ ] very difficult

Comments:

.....

8) Could you please tell us how important are the following factors for a semantic annotation tool?

Seamless integration into other (authoring) applications, e.g. 3ds Max, Photoshop, Cinema 4D…

[ ] very important [ ] important [ ] marginally important [ ] not important

To find again media resources created in previous projects:

[ ] very important [ ] important [ ] marginally important [ ] not important

Possibility to select one of several meanings in case of homonyms (e.g. bank (financial institution) vs. bank(seating))

[ ] very important [ ] important [ ] marginally important [ ] not important

Comments: THIS SEEMED LIKE A LOT OF WORK FOR NOT MUCH RETURN.

9) Could you please tell us how relevant to your business the tools demonstrated is?

[ ] very relevant [ ] relevant [ ] marginally relevant [ ] not relevant

Comments:

.....

10) Please, score the cost impact in your business if using this tool:

[ ] Costs would be decremented
[ ] Costs would be the same
[ ] Costs would be increased
[ ] Costs would be increased, although it would be compensated for a better quality of service

Comments:

.....

12.9 Person 9

7) Was the approach of annotating media resources easy and comprehensible?

[ ] very easy [ ] rather easy [x] rather difficult [ ] very difficult

Comments:

……

8) Could you please tell us how important are the following factors for a semantic annotation tool?

Seamless integration into other (authoring) applications, e.g. 3ds Max, Photoshop, Cinema 4D…

[ ] very important [ ] important [x] marginally important [ ] not important

To find again media resources created in previous projects:

[ ] very important [x] important [ ] marginally important [ ] not important

Possibility to select one of several meanings in case of homonyms (e.g. bank (financial institution) vs. bank(seating))

[ ] very important [x] important [ ] marginally important [ ] not important

Comments:

……

9) Could you please tell us how relevant to your business the tools demonstrated is?

[ ] very relevant [ ] relevant [x] marginally relevant [ ] not relevant

Comments:

……

10) Please, score the cost impact in your business if using this tool:

[ ] Costs would be decremented
[x] Costs would be the same
[ ] Costs would be increased
[ ] Costs would be increased, although it would be compensated for a better quality of service

Comments:

……

12.10 Person 10

7) Was the approach of annotating media resources easy and comprehensible?

[*] very easy [ ] rather easy [ ] rather difficult [ ] very difficult

Comments: Clear and well structured. Will help to save time in the process.

8) Could you please tell us how important are the following factors for a semantic annotation tool?

Seamless integration into other (authoring) applications, e.g. 3ds Max, Photoshop, Cinema 4D…
[*] very important [ ] important [ ] marginally important [ ] not important

To find again media resources created in previous projects:
[ ] very important [*] important [ ] marginally important [ ] not important

Possibility to select one of several meanings in case of homonyms (e.g. bank (financial institution) vs. bank(seating))
[ ] very important [*] important [ ] marginally important [ ] not important

Comments: well organized and detailed data bank is important factor in all projects where previously used data could be reused, by having an easy access to all information needed.

9) Could you please tell us how relevant to your business the tools demonstrated is?

[ ] very relevant [ ] relevant [ ] marginally relevant [*] not relevant

Comments:
……

10) Please, score the cost impact in your business if using this tool:

[ ] Costs would be decremented
[ ] Costs would be the same
[ ] Costs would be increased
[ ] Costs would be increased, although it would be compensated for a better quality of service

Comments:
……

12.11 Person 11

Did not fill out this part of the questionnaire

12.12 Person 12

7) Was the approach of annotating media resources easy and comprehensible?
[ ] very easy [x] rather easy [ ] rather difficult [ ] very difficult

Comments:
Creating the list of materials is being the most precise and time consuming work
……

8) Could you please tell us how important are the following factors for a semantic annotation tool?

Seamless integration into other (authoring) applications, e.g. 3ds Max, Photoshop, Cinema 4D…
[ ] very important [x] important [ ] marginally important [ ] not important

To find again media resources created in previous projects:
[ ] very important [x] important [ ] marginally important [ ] not important

Possibility to select one of several meanings in case of homonyms (e.g. bank (financial institution) vs. bank(seating))

[ ] very important [x] important [ ] marginally important [ ] not important

Comments:

……

9) Could you please tell us how relevant to your business the tools demonstrated is?

[ ] very relevant [ ] relevant [x] marginally relevant [ ] not relevant

Comments:

Probably there would be time when musicians would look for specific material they recorded some time ago… but nonetheless they would not start filing all the material for that purpose.

……

10) Please, score the cost impact in your business if using this tool:

[ ] Costs would be decremented

[ ] Costs would be the same

[x] Costs would be increased

[ ] Costs would be increased, although it would be compensated for a better quality of service

Comments:

Would the material be listed automatically –like, all Shakespeare books scanned in and then it would be easy to look up a specific quote or thought… But to make specific filing a day to day job in order to find a part in years to come – this sounds more like a job for libraries than everyday other business. Too work consuming therefor costly for my use.

……

13 Annex IV: Results of Task Comparison Times

Participant: 1

	Tag Tool			Annotation System			PhotoStuff		
File	**from**	**to**	**time (s)**	**from**	**to**	**time (s)**	**from**	**to**	**time (s)**
1	14:54:20	14:54:45	25	15:03:20	15:04:25	65	15:18:05	15:19:10	65
2	14:55:03	14:55:30	27	15:05:20	15:06:45	85	15:19:20	15:21:20	120
3	14:56:07	14:56:53	46	15:07:25	15:08:40	75	15:21:30	15:23:10	100
4	14:57:36	14:58:00	24	15:09:20	15:10:40	80	15:28:45	15:29:20	35
5	14:58:25	14:58:45	20	15:11:23	15:12:10	47	15:29:30	15:31:00	90
6	14:59:00	14:59:15	15	15:12:30	15:13:00	30	15:31:10	15:32:00	50
7	14:59:45	15:00:03	18	15:13:28	15:14:07	39	15:32:07	15:32:45	38
8	15:00:20	15:00:30	10	15:14:35	15:14:55	20	15:33:00	15:34:10	70
9	15:00:50	15:01:10	20	15:15:20	15:15:45	25	15:34:20	15:35:15	55
10	15:01:47	15:02:05	18	15:16:03	15:16:48	45	15:35:25	15:36:50	85

Participant: 2

	Tag Tool			Annotation System			PhotoStuff		
File	**from**	**to**	**time (s)**	**from**	**to**	**time (s)**	**from**	**to**	**time (s)**
1	14:04:35	14:05:03	28	14:10:00	14:10:25	25	14:20:50	14:22:00	70
2	14:05:15	14:05:30	15	14:10:45	14:11:15	30	14:22:15	14:23:30	75
3	14:05:45	14:06:15	30	14:11:25	14:11:45	20	14:23:43	14:24:45	62
4	14:06:30	14:06:50	20	14:12:10	14:12:45	35	14:24:55	14:25:25	30
5	14:07:01	14:07:18	17	14:13:00	14:13:18	18	14:26:00	14:26:53	53
6	14:07:24	14:07:33	09	14:13:30	14:13:43	13	14:27:00	14:28:00	60
7	14:07:45	14:08:00	15	14:14:00	14:14:15	15	14:28:07	14:28:55	48
8	14:08:06	14:08:25	19	14:14:25	14:14:42	17	14:29:18	14:29:56	38
9	14:08:30	14:08:45	15	14:14:50	14:15:25	35	14:30:15	14:31:05	50
10	14:08:55	14:09:15	20	14:15:40	14:16:15	35	14:31:15	14:31:55	40

Participant: 3

	Tag Tool			Annotation System			PhotoStuff		
File	**from**	**to**	**time (s)**	**from**	**to**	**time (s)**	**from**	**to**	**time (s)**
1	12:03:40	12:04:00	20	11:48:50	11:49:30	40	13:04:10	13:06:00	110
2	12:04:25	12:04:50	25	11:50:30	11:51:30	60	13:06:35	13:09:30	175
3	12:05:07	12:05:35	28	11:52:00	11:53:15	75	13:09:45	13:10:45	60
4	12:05:50	12:06:20	30	11:54:00	11:55:45	105	13:10:55	13:11:25	30
5	12:06:35	12:06:50	15	11:57:00	11:58:00	60	13:11:45	13:14:55	190
6	12:07:05	12:07:12	07	11:58:30	11:58:40	10	13:16:07	13:16:50	43
7	12:07:25	12:07:50	25	11:59:30	12:01:00	90	13:17:05	13:18:00	55
8	12:08:02	12:08:09	07	12:01:25	12:01:30	05	13:18:20	13:20:00	100
9	12:08:25	12:08:32	07	12:02:05	12:02:20	15	13:20:15	13:21:00	45
10	12:08:43	12:08:51	08	12:03:45	12:03:50	05	13:21:25	13:23:30	125

Participant: 4

	Tag Tool			Annotation System			PhotoStuff		
File	**from**	**to**	**time (s)**	**from**	**to**	**time (s)**	**from**	**to**	**time (s)**
1	10:44:30	10:44:50	20	10:52:00	10:52:27	27	11:07:22	11:08:24	62
2	10:45:35	10:45:55	20	10:53:01	10:53:37	36	11:08:54	11:10:48	114
3	10:46:35	10:46:57	22	10:54:16	10:54:55	39	11:11:07	11:12:00	53
4	10:47:15	10:47:34	19	10:55:22	10:56:16	54	11:12:25	11:12:53	28
5	10:48:00	10:48:20	20	10:56:51	10:57:10	19	11:13:10	11:14:46	96

6	10:48:33	10:48:41	08	10:57:31	10:57:37	06	11:15:02	11:15:48	46
7	10:49:11	10:49:25	14	10:58:20	10:58:34	14	11:16:00	11:16:35	35
8	10:49:41	10:49:56	15	10:58:55	10:59:07	12	11:16:46	11:18:07	81
9	10:50:17	10:50:26	09	10:59:32	10:59:46	14	11:18:23	11:19:08	45
10	10:50:41	10:50:53	12	11:00:05	11:00:17	12	11:19:18	11:20:25	67