An Intelligent Media Framework for Multimedia Content

Tobias Bürger

Salzburg Research, Jakob-Haringer-Str. 5/III, 5020 Salzburg, Austria,
`tobias.buerger@salzburgresearch.at`,
WWW home page: `http://www.salzburgresearch.at/~tbuerger`

Abstract. Search, retrieval and navigation in multimedia repositories is a task common to all multimedia management systems: Users are supported by a wide range of features which are traditionally based on full text search and metadata queries. However generating metadata is an error-prune and work-intensive task, that for multimedia content cannot yet be made fully automatically. In this position paper we describe our vision of an Intelligent Media Framework that is capable of combining metadata and knowledge about media items in order to support user-orientation, search and retrieval in media-rich information spaces: We try to integrate heterogeneous sources to create an Intelligent Media Framework containing Intelligent Media Objects carrying behavioural knowledge and capable of fully describing themselves. The properties of these objects amongst others serve to the fact that users more likely search by the "meaning" of audiovisual objects and what is represented by them respectively, than by their pure low-level features.

1 Introduction

As we have experienced in a former project [5], semantic descriptions of content can enhance fast and easy navigation through audio-visual repositories. Semantics or an interpretation of the content is important to make content machine-processable and to make it even useful in goal-oriented workflows that shall support knowledge workers in fulfilling certain tasks. Some of the recent research projects in for eg. the area of semantic (or symbolic) video annotation try to derive the semantics from the videos' low level features or from any other available basic metadata. Some - that rely on highly automated indexing - build upon the results of automatic speech recognition however error-prone they are. Most of these approaches are - as also pointed out in [3] - not capable of fully exploiting the semantics of multimedia content. Therefore we propose an approach that tries to gather metadata from diverse sources to form the basis for Intelligent Media Objects (IMOs) that amongst others are able to describe themselves, can be traded or exchanged. Those IMOs are the core of our vision of an Intelligent Media Framework (IMF).

The remaining paper is organised as follows: In section 2 IMOs are introduced, section 3 contains design considerations for an environment based on those IMOs and section 4 concludes with a description of our ongoing and future work.

2   A Model for Intelligent Media Objects

Intelligent Media Objects (IMOs) are objects that integrate metadata and cognition based knowledge about media items that are included or referred by them. Additionally they carry explicit semantic descriptions of their properties that enable to communicate certain aspects - like how to use, transfer or change them - to foreign systems.

IMOs - as we understand them - are inspired by the vision of *"smart content objects"* that were first introduced in the EP2010 study, which was a strategic European study on the future of electronic publishing towards 2010 [2].

The following definition of *"smart content"* is adopted from Steven R. Newcomb who defined it by its two key properties:

- *Smart content is formally self-describing, and*
- *smart content is interpretable in such a way as to support semantic indexing, so that master indexes can be automatically generated from diverse, arbitrary combinations of smart sources, ad libitum [9].*
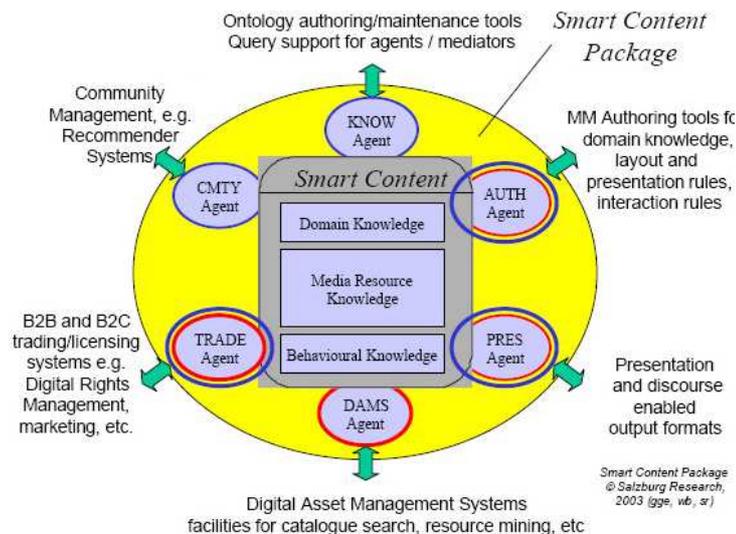

Fig. 1. A *"Smart Content"* package, ©Salzburg Research

As shown in figure 1, a *"smart content"* - package primarily carries content and behavioural knowledge and has points of contact to some ambient environment in which the main external actors are designed to understand *"smart content objects"*. These contact points can be viewed as service interfaces that can be used to query properties or act on data [2].

A realization of the *"smart content"* - vision, which can be seen as an abstract foundation for IMOs, are the so called knowledge content objects (KCOs), that were developed in the recently finished IST project METOKIS[1]: In METOKIS a model to describe knowledge enhanced multimedia content objects was developed [1]. The knowledge in KCOs is modeled by using three levels of knowledge:

1. It includes **content** which means the actual media file (eg. a file, a stream or a text),
2. **knowledge about the content** which means metadata and knowledge specific to the content object (e.g. the encoding/format or colour coding scheme), and
3. **knowledge about the subject of the content** (the topic) which means knowledge about the meaning of the content, a conceptualization or a description that explains what the content is about.
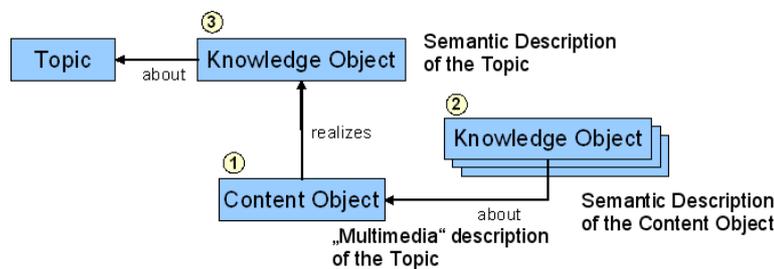

Fig. 2. The Structure of Knowledge Content Objects (KCO), ©Salzburg Research

As shown in figure 2, a content object and several knowledge objects are the main parts of a KCO. There are two types of knowledge objects according to the three levels of knowledge defined above:

1. The first type is knowledge which is bound to the specific content object that describes different aspects of the content and
2. the second type of knowledge object represents semantic descriptions about a topic (or subject), that can be realized by multiple content objects.

The different knowledge objects are structured in so-called semantic facets that are modular interfaces to describe the properties of KCOs, which can be seen as the counterpart to the service interfaces of *"smart content objects"*, that were described above.

The knowledge structure of KCOs as presented above is based on the Information Object design pattern that is defined in an extension of the DOLCE foundational ontology [7, 1].

---

[1] http://metokis.salzburgresearch.at

A more detailed description of the structure and semantics of KCOs can be found in [1].

To sum up: what exactly are or will IMOs be? IMOs are *"smart content objects"*, which are realized through KCOs and which will be instantiated for an application domain dealing with audiovisual content.

3 An Intelligent Media Framework for Multimedia Content

The Intelligent Media Framework (IMF) has to support and recommend multimedia objects (eg. video streams, animation films or presentations) to knowledge workers, in order to support easy and fast navigation in vast information spaces. Intelligent Media Objects (IMOs) - as presented in section 2 - are the integral part of this framework. To model IMOs for the IMF we have to - amongst others - get a deep understanding of the domain where these objects will be based in. Further we have to model the media included and the targeted end-users that will handle them. Thus we will have to develop domain models and user models and integrate them with standardized (multi-) media models, all that with respect to ongoing standardization activities. For that task, a deep understanding of the relation of multimedia models (e.g. MPEG-7 [6], MPEG-21 [4]) with upper level ontologies (i.e. domain ontologies) has to be developed (see for example [8]).

The intention of IMOs in the IMF is to map the knowledge included in them with personal information of users in order to support their contextualized delivery. The properties of IMOs also allow to define knowledge objects beforehand and afterwards append content to it, which means to find content that suits to the topics that are modeled in the knowledge object.

For all that, it is necessary to understand what the multimedia objects are about. However, as it is not possible to automatically annotate and derive the semantics of audiovisual material without human intervention, we envisage metadata and background knowledge to come from diverse sources: First of all, knowledge will be automatically detected and extracted from media data, secondly knowledge and content from external resources will be automatically integrated and thirdly semi-automatic annotation by users will be supported. Trying to integrate these reliable external resources instead of purely relying on automatic analysis means going one step back from recent efforts, but will surely produce much better results, which are important for productive (real-time) scenarios.

4 Ongoing and Future Work

Salzburg Research participated in several national or European projects that dealt with the concept of intelligent content. Two of them were the IST projects CULTOS[2] and METOKIS, that developed models for intelligent content objects

---

[2] http://www.cultos.org

which our group in the Salzburg New Media Lab[3] tries to adapt in the currently running project Smart Content Factory[4] and the recently started IST project LIVE[5]. In LIVE our main responsibility lies in the definition of an intelligent media framework to support broadcasters in the live staging of media events. For LIVE we will develop a specialization of knowledge content objects (KCOs), which were developed in METOKIS, for the domain of staging live events.

References

1. Behrendt, W., Gangemi, A., Maass, W., Westenthaler, R.: Towards an Ontology-Based Distributed Architecture for Paid Content In: A. Gomez-Perez and J. Euzenat (Eds.) : ESWC 2005, LNCS 3532, pp. 257-271, 2005.
2. Behrendt, W., Geser, G., Mulrenin, A.: EP2010 - the Future of Electronic Publishing Towards 2010, 2003.
3. Bloehdorn, S. et al.: Semantic Annotation of Images and Videos for Multimedia Analysis In Proceedings of the 2nd European Semantic Web Conference, ESWC 2005, Heraklion, Greece, May 2005.
4. Bormans, J. and Hill, K.: Mpeg-21 overview v.5. http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm, 2002. ISO/IEC JTC1/SC29/WG11. (last visited: 06.04.2006)
5. Bürger, T.,Gams, E., Güntner, G.: Smart Content Factory - Assisting Search for Digital Objects by Generic Linking Concepts to Multimedia Content In: Proceedings of the Sixteenth ACM Conference on Hypertext and Hypermedia (HT '05), 2005.
6. Martinez-Sanchez, J. M., Koenen, R. and Pereira, F.: Mpeg-7: The generic multimedia content description standard, part 1. IEEE MultiMedia, 9(2): pp. 78-87, 2002.
7. Masolo, C. et al.: The WonderWeb Library of Foundational Ontologies. WonderWeb Deliverable D17 - Preliminary Report (ver. 29-05-2003), 2003. http://www.loa-cnr.it/Papers/DOLCE2.1-FOL.pdf (last visited: 06.04.2006)
8. Nack, F., van Ossenbruggen, J., Hardman, L.: That Obscure Object of Desire: Multimedia Metadata on the Web, Part II. IEEE Multimedia, 12(1):54-63, January-March 2005.
9. Newcomb, S. and Blunden, B.: Smart Content, Opportunity Discovery, Knowledge Trading and Alliance Making (slides), presented at XML Europe 2004, Amsterdam, April 21, 2004.

---

[3] http://www.newmedialab.at

[4] http://scf.salzburgresearch.at

[5] http://www.ist-live.org