

The Need for Formalizing Media Semantics in the Games and Entertainment Industry

Tobias Bürger

(Digital Enterprise Research Institute (DERI), Innsbruck, Austria
tobias.buerger@deri.at)

Herwig Zeiner

(Joanneum Research, Graz, Austria
herwig.zeiner@joanneum.at)

Abstract: The digital media and games industry is one of the biggest IT based industries worldwide. Recent observations therein showed that current production workflows may be potentially improved as multimedia objects are mostly created from scratch due to insufficient reusability capacities of existing tools. In this paper we provide reasons for that, provide a potential solution based on semantic technologies, show the potential of ontologies, and provide scenarios for the application of semantic technologies in the digital media and games industry.

Key Words: multimedia semantics, digital entertainment, ontologies, semantic search
Category: H.5.1

1 Introduction

As initial investigations in the SALERO¹ project showed, most multimedia objects are created from scratch due to insufficient reusability capabilities of existing tools. One reason for that is their poor self-description-ability and the lack of formal representations of the properties of multimedia objects, their context and intended meaning. In order to reach a sufficient solution for that problem, several steps need to be solved: Besides the need for algorithms and frameworks to automatically extract high-level semantics from low-level features² which is well known as the "Semantic Gap" [Smeulders et al. 2000], annotation support for users, a solution for metadata interoperability across the content lifecycle, and cross-media adaptation is needed.

In this paper, we first introduce the intention of (multimedia) ontologies, its purpose and potential benefits for media production, sketch scenarios in which they can be applied, and finally conclude with an outlook.

¹ The overall goal of the integrated project SALERO (<http://www.salero.info>) is to define and develop intelligent content with context-aware behaviours for self-adaptive use and delivery across different platforms, building on and extending research in media technologies, web semantics to reverse the trend toward ever-increasing cost of creating media

² eg. "this picture depicts a scene in a football game" is inferred from the low level features "white circle AND green background color"

2 Using Ontologies for Semantic Representation of Media Items

As defined by Gruber [Gruber 1993], an ontology is an explicit specification of a (shared) conceptualization. The term ontology has been in use for many centuries and ontologies are widely used in applications related to information integration, information retrieval, knowledge management or in the Semantic Web. Ontologies are usually used to establish a common understanding of a domain and to capture the domain knowledge. This is usually done by modeling basic terms and relations which hold between terms, and by providing rules stating restrictions on the usage of both terms and relations.

In SALERO, we try to establish a **multimedia ontology framework** that combines declarative descriptions of

1. *Low-level physical and semantic features* through the use of multimedia description standards like MPEG-7 [Martinez et al. 2002] or essence internal formats
2. Domain specific *high-level semantic features* through the use of ontology languages like WSML [de Bruijn et al. 2005] or OWL [Dean et al. 2006]
3. *context information and rules* using WSML or RIF [Boley and Kifer 2007]

By using multimedia ontologies, recent research initiatives in the multimedia domain try to overcome the commonly known drawbacks of existing multimedia metadata standards for the descriptions of the semantics of multimedia content (see e.g. [Bloehdorn et al. 2005, Troncy et al. 2006, Benitez et al. 2002]).

Multimedia ontologies are mostly designed to serve one or more of the following purposes [Eleftherohorinou et al. 2006]:

- *Annotation*, which is in most cases motivated by the need to have high-level summarizations of the content of multimedia items
- *Automated semantic analysis*, i.e. to support the analysis of the semantics and syntax of the structure and content of multimedia items
- *Retrieval*, i.e. to use rich formal descriptions to enable context-based retrieval and recommendations to users. The use of semantics enables automatic matching of content properties with user properties
- *Reasoning*, i.e. the application of reasoning techniques to discover previously unknown facts of multimedia content or to enable question answering about properties of the content.

- *Personalized filtering*, ie. the delivery of multimedia content according to user-, network- or device-preferences.
- *Meta-Modeling*, ie. to use ontologies or rules to model multimedia items and associated processes.

3 The Purpose of Multimedia Ontologies in the digital games and entertainment industries

The potential benefits of formalizing media semantics were summarized and highlighted already before [van Ossenbruggen et al. 2004, Nack et al.. 2005]. In order to highlight the benefits for the digital games and entertainment industries, we especially try to summarize the purpose of the ontology framework that will be built in SALERO, sketch expected benefits for traditional media production and point out to important problems that we see in each point.

3.1 Semantic Search

To enable semantic search is the prime advantage of using ontologies in media production. Semantic Search aims to improve recall and precision of search results for multimedia objects. A prime pre-requisite for this is a way to attach hidden or contextual features to media items which are not visually embedded in them. By encoding such information using ontologies, the ontologies can then be used for the retrieval process and to present the results. This helps to clearer present contextual information, and helps to find more accurate results. A semantic search facility also provides a high-level means to explore collections with a high precision and recall.

The *most important problems* that we intend to solve with this facility are

1. The *Semantic Gap*, i.e. how to assign meta-data (semi-)automatically to multimedia data?
2. Low Precision/Recall for search in large multimedia collections or how to increase the amount of true positives for multimedia information retrieval?
3. How to match context of data with context of users?

3.2 Annotation Support

Mmetadata³ has an important role in the multimedia lifecycle which was already highlighted by many others [Smith and Schirling 2006]. A recurring problem in

³ metadata is data about data

industrial settings is however the task of creating metadata and keeping it up to date. Therefore one of the biggest issues we see is how to support creative people in creating annotations and how in turn existing workflows are only minimally disturbed by the implementation of new annotation facilities. We aim to develop ontology-based ways of supporting the user with these important tasks by developing a suite of ontology tools which can be used in daily work. An important point – as already emphasized above – is the integration of the ontology-enhancements into the work process with a minimum amount of disturbing the workflow. We intend to solve this by providing APIs that allow to include the functionality into existing multimedia authoring tools.

The *most important problems* that we intend to solve with this facility are

1. How can creative – and sometimes non-technical – people create ontological annotations?
2. How to maximize support and simultaneously minimize disturbance of current production workflows?

3.3 Unifying Disparate Metadata Formats

At present, different metadata standards are used to annotate in- and output of different steps in the multimedia production lifecycle. This fact is highlighted for the broadcasting domain in [Smith and Schirling 2006]. One result of SALERO will be to show the feasibility of using ontologies to unify these partially disparate metadata formats and the vocabularies used in them. Using ontologies is promising because of their modeling power, their formal background and strong their semantics compared to ad-hoc and informally specified data models. We try to use the arising multimedia ontology in order to enhance metadata integration and as a further step to improve the reusability of multimedia items.

The *most important problems* that we intend to solve with this facility are

1. How to process and re-use assets in different production tools?
2. How to integrate different standards that are used to describe different asset types? (eg. to provide a unifying search infrastructure on top of asset repositories)
3. How to create a scalable mapping/mediation layer between each of the standards? (ie. a bidirectional mapping between each of the standards does not scale!)

3.4 Support for Cross Media Adaptation

Another cost driver of today's media productions is the huge manual effort necessary for the adaptation of media to different target platforms and output formats (e.g. cinema, games, print or Internet). Ontologies offer the possibility to model and capture a rich set of metadata including the context of images or other assets. Therefore they can be used to cross purpose multimedia assets as automatically as possible. A wide range of multimedia objects is used by different parties in different media productions. This has to be considered when choosing or developing applications to create, manage or use ontologies for description of multimedia data like image sequences (TV recordings, computer generated content, film sequences), audio objects (recorded sound as well as synthesized speech), 3D Objects (Animated 3D objects in the application domains of interactive games, special effects for film), or any combination of the mentioned types.

The *most important problems* that we intend to solve with this facility are

1. How to ensure a common technological basis for (originally diverse) production tools?
2. How to establish declarative descriptions of workflows in order to match process- with content-descriptions for Cross-Media-Production?

4 Scenarios - Using Multimedia Ontologies in Media Production

In this section we briefly sketch examples how a multimedia ontology framework may be applied in media production environments.

4.1 Ontologies to Support the Re-Use of Assets

One big problem that media production companies are facing is the lack of possibilities to re-use material across productions. This is mainly due to the lack of definitions of methods and rules how assets can be reused and how certain assets can adapt themselves to new environments, eg. it needs to be considered how characters are able to interact with the re-used elements or how the elements can be adjusted to fit in a particular scene. In order to recognize if assets are re-useable their properties have to be stored explicitly (ie. the usage context of the asset, the rights to re-use, or rules about how to extract specific parts of an asset in order to be transferable to other animations. Using ontologies to describe the (usage) context of assets could help to clearly identify which methods are needed to transfer assets between different productions and to automatically identify the parts of the animations that one wants to re-use. This demands

for advanced asset management systems that are able to store rich metadata together with the assets, perform fast and reliable searches and to access assets across productions.

4.2 Integrated Production and Cross-Media Delivery of Assets

In the increasingly fragmented media distribution marketplace there is a great need to be able to produce different delivery formats in parallel. One example for that is the sharing of datasets between film asset creation and tie in-game production. There are different problems attached to that wish: First of all there is the problem of reusing media objects in new functions: e.g. media objects in films and games have different roles and functions: in films they are actors, in games they are avatars: in films, they can be seen and heard, in games they are used in various interactive ways. This change of function affects the design of media objects. This demands for explicit descriptions of the story, storytelling, style and medium of a production. There ontologies could help. In turn, rules apply on how to convert between different genres. These rules could be modelled using ontology languages in order to turn them into actionable knowledge that makes an automatic conversion possible.

4.3 Ontologies to Aid Multimedia Information Retrieval

A major research problem in multimedia information retrieval is the "Semantic Gap" [Smeulders et al. 2000], the large gulf between the low level image features which can typically be processed in a multimedia document, and the high level concepts which a user is typically interested in. For example, a user may want to search for a video showing "Bing and Bong on a trip to planet XY"⁴. The high level concepts implicit in this query may be stated as the characters "Bing" and "Bong", plus the abstract action "traveling". The action "traveling" in this case is almost impossible to recognize and could be derived from other recognizable features like "sitting on a couch"⁵ or "galaxy"⁶ which has to be explicitly modeled in an ontology capturing the knowledge of this special application domain. Automatic annotation of videos and images is currently an active research topic, and allows data-driven techniques to be used together with large training sets. However, automatic processing is not sufficient to annotate multimedia documents with the features like the ones used in the example above. Therefore annotation tools need to be developed that allow designers to annotate the material during the production to ease a latter retrieval-task.

⁴ Bing and Bong virtual characters and part of Tinyplanets (see <http://www.tinyplanets.com>) which is a UK television show aimed at pre-schoolers

⁵ This is Bing and Bong's favorite vehicle

⁶ Bing and Bong usually travel in the galaxy

4.4 Ontologies as an Aid to Personalising Search

Search tasks, such as those supported by Multimedia Information Retrieval (MIR) systems, are typically subtasks of some main work task, which may be the creation of a new cartoon character, or some other aspect of the user's work. As such there are many contextual factors which may be captured about users, encoded in an ontology, with the aim of providing better retrieval results for the user in that particular situation. For example, the role of a user within an organisation may imply different search preferences: managers and administration staff may be less technical minded than engineering staff, and therefore an information retrieval engine should attempt to find documents of a less technical difficulty than may be presented to an engineer or scientist. Likewise, the immediate work context of the user engaged on the creation of characters for a particular film, may be used to condition searches, with the aim of interpreting queries relative to this topic.

5 Conclusions

In general, formal semantics can support the annotation, analysis, retrieval or reasoning about multimedia assets. With this paper we aimed to stress the importance of the use of formal semantics in the digital games and entertainment industry by pointing out to expected benefits and by sketching scenarios illustrating their intended application in media production in general and the SALERO project in particular.

Our next steps in SALERO are to engineer the first version of the ontologies and implement a first version of the ontology management framework to support the different media production scenarios in the SALERO project.

Acknowledgements

The research leading to this paper was partially supported by the European Commission under contract IST-FP6-027122 "SALERO".

References

- [Benitez et al. 2002] Benitez, A. B., Rising, H., Jörgensen, C., Leonardi, R., Bugatti, A., Hasida, K., Mehrotra, R., Tekalp, A. M., Ekin, A., and Walker, T.: "Semantics of Multimedia in Mpeg-7"; Proceedings of the IEEE International Conference on Image Processing (ICIP 2002), pp. 137–140, 2002.
- [Bloehdorn et al. 2005] Stephan Blöhdorn, Kosmas Petridis, Carsten Saathoff, Nikos Simou, Vassilis Tzouvaras, Yannis Avrithis, Siegfried Handschuh, Yiannis Kompatsiaris, Steffen Staab, and Michael G. Strintzis: "Semantic Annotation of Images and Videos for Multimedia Analysis"; Proceedings of the Second European Semantic Web Conference, ESWC 2005, Heraklion, Greece, pp. 592–607. Springer, May 2005.

- [Boley and Kifer 2007] Harold Boley and Michael Kifer (eds.): "RIF Core Design", W3C Working Draft 30 March 2007, 2007. available online: <http://www.w3.org/TR/rif-core/> (last accessed: 01.04.2007)
- [de Bruijn et al. 2005] de Bruijn (ed.): "The Web Service Modeling Language WSML"; WSML Final Draft, 20 March 2005, Deliverable D16.1v0.2, 2005. Available online: <http://www.wsmo.org/TR/d16/d16.1/v0.2/> (last accessed: 01.04.2007)
- [Buerger and Westenthaler 2006] Bürger Tobias and Rupert Westenthaler: "Mind the Gap - Requirements for the Combination of Content and Knowledge"; Proceedings of the first international conference on Semantics And digital Media Technology (SAMT), December 6-8, 2006, Athens, Greece.
- [Dean et al. 2006] Dean M. and Schreiber G (eds.): "OWL Web Ontology Language Reference"; W3C Recommendation, 2006. Available online: <http://www.w3.org/TR/owl-ref/> (last accessed: 01.04.2007)
- [Eleftherohorinou et al. 2006] Hariklia Eleftherohorinou, Vasiliki Zervaki, Anastasios Gounaris, Vasileios Papastathis, Yiannis Kompatsiaris, and Paola Hobson: "Towards a Common Multimedia Ontology Framework"; Analysis of the Contributions to Call for a Common multimedia Ontology Framework Requirement; April, 2006. Available online: http://www.acemedia.org/aceMedia/files/multimedia_ontology/cfr/MM-Ontologies-Reqs-v1.3.pdf (last accessed: 01.04.2007)
- [Gruber 1993] Gruber, T. R.: "A translation approach to portable ontology specifications"; Knowledge Acquisition, vol. 5, no. 2, pp. 199–220, Juni, 1993.
- [Martinez et al. 2002] Martinez-Sanchez, J. M., Koenen, R., and Pereira, F.: "Mpeg-7: The generic mul- timedia content description standard, part 1"; IEEE MultiMedia, vol. 9, no. 2, pp 78–87, 2002.
- [Nack et al. 2005] Nack, F., van Ossenbruggen, J., and Hardman, L.: "That Obscure Object of Desire: Multimedia Metadata on the Web"; Part 2 IEEE MultiMedia, IEEE Computer Society, 2005, 12, 54-63
- [van Ossenbruggen et al. 2004] van Ossenbruggen, J., Nack, F., and Hardman, L.: "That Obscure Object of Desire: Multimedia Metadata on the Web"; Part 1 IEEE MultiMedia, IEEE Computer Society, 2004, 11, 38-48
- [Smeulders et al. 2000] Arnold W.M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, Ramesh Jain: "Content-Based Image Retrieval at the End of the Early Years"; IEEE Transactions on Pattern Analysis and Machine Intelligence ,vol. 22, no. 12, pp. 1349–1380, December, 2000.
- [Smith and Schirling 2006] John R. Smith and Peter Schirling: "Metadata standards roundup"; IEEE Multimedia, vol. 13, no. 1, 2006, pp. 84–88, 2006.
- [Troncy et al. 2006] R. Troncy, W. Bailer, M. Hausenblas, P. Hofmair, R. Schlatte.: "Enabling Multimedia Metadata Interoperability by Defining Formal Semantics of MPEG-7 Profiles"; Proceedings of the 1st International Conference on Semantics And Digital Media Technology (SAMT 06), Athens, Greece, December 2006.