DMS-1 driven Data Model to enable a Semantic Middleware for Multimedia Information Retrieval in a Broadcaster

Gorka Marcos, Kevin Alonso, Igor G. Olaizola, Julián Flórez.
Vicomtech, Department of Digital TV & Multimedia Services
Mikeletegi Pasealekua, 2009 San Sebastian , Spain
gmarcos,kalonso,iolaizola,jflorez@vicomtech.org

Arantza Illarramendi
LSI Department, University of the Basque Country, Apdo. 649, 20080 Donostia-San Sebastián, Spain
a.illarramendi@ehu.es

Abstract

This article presents the motivation and the implementation of a semantic model developed to support diverse semantic services in a Multimedia Asset Management system in a Broadcaster. The model is mainly driven by DMS-1 (Descriptive Metadata Scheme) standard, which is part of the Multimedia Exchange Format standard defined by the broadcast industrial community and according to our knowledge we propose the first implementation of it using the OWL language. This model has been complemented with other models coming from the academia in order to cover the diverse nature of the different semantic needs identified in the whole workflow.

1. Introduction

The multimedia information management and retrieval has been a very active field. In the professional context, the digitalization has lead to a great revolution in the way the companies handle, store and make their content accessible. Our work is contextualized in the Broadcast sector, where the tapes are disappearing and the workflow and the access to the information have become concurrent and, in many cases, immediate.[1]

Apart from the new opportunities brought by these changes [15], the multimedia processing techniques and the technologies coming from the semantic web [26] have provided new ways to optimize the management of the digital assets. In the next section we will mention some important contribution on very diverse topics that are related with different subprocesses of the workflow.

In this context, this article presents a semantic model designed to provide those services that rely on the semantics of the assets while keeping the perspective of the broadcast industry. In order to do that, the model has been built on the top of an ontology that partially implements the DMS-1 [24] (Descriptive Mtadata Scheme) of the MXF [11] (Multimedia eXchange Format), which is a standard widely adopted by that industry.

This article is structured according to as follows. The second section provides a brief overview of the related work in multimedia modeling. The overview emphasizes the work performed in the broadcast environment. In the following section a short introduction the the DMS-1 standard is included. In the fourth section, the context of the model and its implementation are described. The work ends highlighting some conclusions and with some thoughts about the future work.

2 Related Work

From both the industrial and the academic communities, very important efforts have been made in order to standardize the definition and management of metadata in the broadcast environment. In [24, 6] the authors provide an overview of the main standards. Some of the most remarkable initiatives are the SMPTE descriptive metadata dictionaries and their evolution into the DMS-1 [24] of MXF [11], the EBU P/Meta standard for annotation, the MPEG-7 standard [27],Dublin Core [5], TVAnytime [19] and the Broadcast Metadata exchange Format (BMF) [12].

Furthermore, there is another very active research field which is not focused on the broadcast context but in the reasoning and retrieval of multimedia items applying semantic data models and ontologies to improve some of the processes of such retrieval. Exhaustive literature has been

presented dealing with the use of semantic techniques to improve the query processing and natural language processing [18], query expansion [4], query adaptation and federation [13], content analysis [23, 25], information integration [13], results ranking [22] and information visualization [7].

MPEG-7 represents the clearest linkage between the standardization efforts in the broadcast environment and the approaches that apply semantic techniques to the different steps of the multimedia information retrieval. A clear example of this is the work of Schallauer et al. [20] work where the reader can find a MPEG-7 based metadata description infrastructure for an audiovisual media processing system including content-analysis, documentation, search and retrieval functionalities. They propose some formalization of the semantic constraints in the MPEG-7 profiles to enable interoperability and automatic use for MPEG-7 based applications. However, despite the reduction of the complexity of using this standard, the presence of solutions or products based on this standard is very low [10]. Furthermore, while some cameras and ingesting devices are handling MXF/DMS-1 based metadata, the emerging techniques to improve semantic driven multimedia analysis [25, 8] are based on customized or MPEG-7 ontologies.

Taking into account this scenario, in order to enable the integration of such emerging techniques, we propose in this paper a metadata model that integrates an ontology of the DMS-1 standard which is interrelated with other ontologies, including an MPEG-7 one, to satisfy all the semantic needs of the retrieval workflow.

The main contribution behind this is that this model is designed to allow as seamless as possible the integration of the emerging semantic techniques that are being applied along the multimedia workflow while keeping the perspective of an industrial standard.

3 DMS-1 in MXF

MXF Descriptive Metadata Scheme 1 (DMS-1) [28], is composed by a set of Descriptive Metadata Frameworks. Each of them is defined by a logical structure of metadata sets that allows them to be used as a plug-in to the Header Metadata of a Material Exchange Format (MXF) file. This provides external access to the metadata related to the different semantic levels represented by each one of the frameworks. According to the standard,

- **Production Framework:** Compiles the set of metadata related to the identification and ownership details of the audio-visual content. MXF considers that during the Production, the metadata is always related to the complete MXF file. Thus, this framework does not deal with the annotation of partial segments. Some of the metadata sets included in the Production Framework are: Publicator, Awards, Classification, Contract, Image Format, Caption description, Episodic Item, Seetting Period and Rights.

- **Clip Framework:** Handles the sets of descriptive metadata and the properties related to the caption and creation of information of the independent audio-visual clips. In MXF a clip is an essence container that may comprise of a number of interleaved audio, video, or data essence elements. Some of the metadata sets covered by this framework are: Shot, Device Parameter, Rights, Classification, Scripting, Processing and Project.

- **Scene Framework:** Contains descriptive metadata sets and properties that describe actions and events within individual scenes of the audio-visual content. Scenes may overlap and they may relate to a point in time rather than having duration.Some of the metadata sets are Setting Period, Annotation, Person, Organisation, Location, Shot, and Titles.

The combination of the above logical frameworks (production, scene and clip) constitutes a powerful and shared descriptive annotation model that has been adopted by manufacturers of professional equipment over the whole workflow.

4 DMS-1 driven metadata model

This section is devoted to the description of the metadata model that we present in this article. It introduces the context that leaded the development of the DMS-1 driven metadata model. The understanding of this context will support the reading of the following subsections that detail its implementation.

4.1 Context of the model

The metadata model that we describe in this paper covers just the passive component of a module that provides different semantic services in a multimedia search and retrieval system. This system has been developed in the RUSHES project[1] [21].

As the reader may see in Fig. 1, the Metadata Model interacts with different modules of the RUSHES system. Its main role is the provision of any service required by those modules that could imply the interaction with the semantic model that we present here. For instance, some of these services are related with the support for the automatic annotation modules (e.g. ontology based reasoning, information preservation), recommendation module, visualization tools and others.

[1] RUSHES project(FP6-045189)website. http://www.rushes-project.eu/

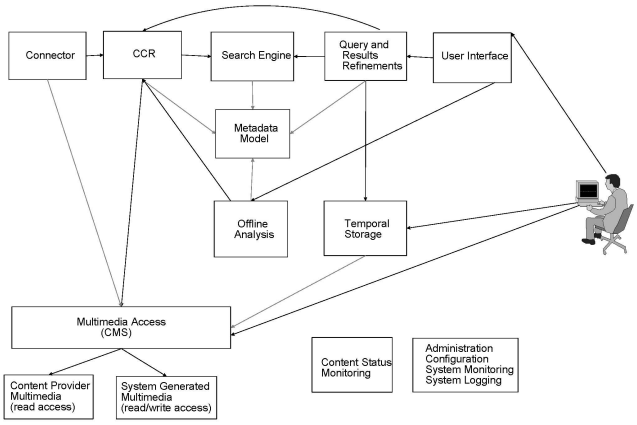

Figure 1. General architecture of the RUSHES system

The model itself has also been deeply influenced by the fact that it had to be used to drive the information exchange between the modules of the system according to the DMS-1 descriptive scheme.

In [21] the reader may find further information regarding the RUSHES system, and in [16] a deeper explanation of the metadata model module is available. As we have previously stated, in this paper we will only describe the passive component of the module. As it will be clarified in the next section, it is mainly built by merging different ontologies, having a DMS-1 ontology as the master ontology.

4.2 Implementation of the model

In the RUSHES system, the model has been implemented in OWL [9] using the Protégé OWL editor [14]. The model itself was composed by three different ontologies. In order to facilitate their maintenance the different ontologies were kept as independent entities and imported by a simple central ontology.

The three ontologies were related with different aspects of the domain of the application. First of all an ontology that descriptively modeled the multimedia items of the broadcaster following the DMS-1 standard. The second ontology was related with the domain covered by the content itself (i.e. news, sports...). And finally, an ontology that covered the descriptive and technical decomposition of the multimedia items.

4.2.1 Descriptive Metadata Scheme (DMS-1) Ontology

The motivation behind this ontology was to implement the descriptive structure of the multimedia assets and the metadata about them. In Fig. 2 the reader may see a partial view o our OWL implementation of the standard in RUSHES.

As has been stated before, the standard groups the annotations according to three logical entities: the production perspective of the asset, the asset as a clip or continuous essence element or element interval and, finally, the different scenes or ”actions and events within individual parts of the audiovisual content”. For each one of the scenes, different metadata sets of properties are defined. For each of them, the cardinality and mandatory level are provided.

The sets of metadata and properties can be applied to various frameworks (Production, Clip and Scene) and their nature is diverse. For example the standard provides sets of metadata about the titles of the asset, awards, events, information about the device parameters, file, formats and rights among many others.

Regarding the temporal decomposition, the ontology, accordingly to the standard, models also the concept ”Shot”. This entity is related with some of the frameworks and allows to define concrete annotations for a shot of any of the essences of the asset (video, audio, and so on). For each of the shot, the duration, start position, and description are provided. And if needed, keypoints or some extra annotations can be attached to the shot.

In order to add the concrete metadata, the ontology models as a concept what the standard defines as the property set named ”Annotation”. This concept is directly linked to different thesauri. This makes the model a powerful tool to capture the metadata for the different pieces of the asset.

4.2.2 Domain Ontology

In the RUSHES system, the domain of information was defined by the raw content repository of the project. This repository was created according to a concrete set of scenarios defined by the broadcaster. These scenarios and the content were very related with the daily activity of the preparation of the news.

The ontology developed in the project was an extension of a well known ontology about news: the LSCOM Lite ontology proposed by Neo et al. [17]. The extension of the ontology was driven by all the information that could be detected by the analysis modules and by the thesauri that are provided to the journalists in the manual annotation tool.

4.2.3 Low Level Ontology

In order to support the conduction of the information during the analysis process, the MDM models the multimedia assets from the video analysis theory perspective. Therefore, the model has to represent the low-level features (histograms, descriptors, etc.

Due to the current requirements of the RUSHES system, our work only tackles the visual description of the content.

In order to achieve this, the MDM incorporates an OWL implementation of the detailed A/V profile (DAVP) proposed in [2, 3].

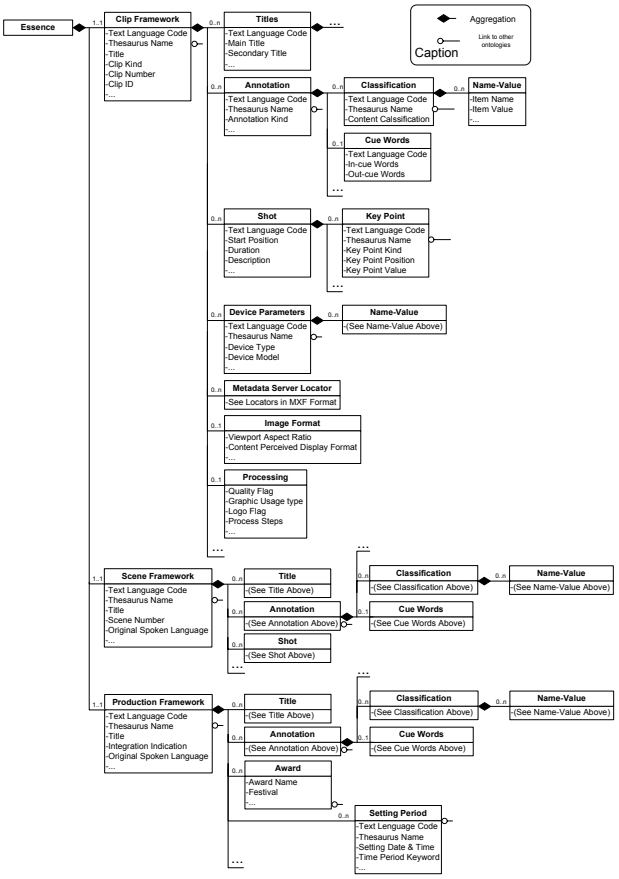

Figure 2. Fragment of the implemented DMS-1 ontology based on the aggregation relation

4.2.4 Merging of the ontologies

The mapping between the DMS-1 ontology and the domain model has been realized by linking the dictionaries proposed by the SMPTE 380:DMS-1 and the concepts of the domain ontology.

The linkage of DAVP ontology with the DMS-1 ontology, has been done through the temporal decomposition of the profile, where the two ontologies intersect.

4.3 Benefits of the usage of the model

The presented model enables the following functionalities in RUSHES.

DMS-1 driven information management This issue covers different parts of the workflow. For instance, the system captures the DMS-1 metadata provided by the Panasonic P2 cameras in a very straightforward and information lossless way. Beside this, the information exchange between the different modules is performed following an XML schema compliant with the DMS-1 standard. The metadata model is able to export the semantic information gathered in the ontology (including the individuals) in a coherent way with that XML.

Information Conduction during the analysis process The inclusion of the ontology that covers the visual descriptors of the MPEG-7 allows the metadata model component to ensure that the information related with the different analysis modules is preserved. Beside this, the model is able to play a key role during the process of the content analysis, since it is able to provide any module with any piece of information that could have have been extracted by analysis module already performed (i.e. list of keyframes provided by the shot boundary detector).

Enabling the reasoning The inclusion of the domain ontology provides support for the information gathering during the analysis process, which enables the ontology based reasoning (e.g. the location of a shot with the presence of a helicopter is classified as outdoor).

Supporting the manual enrichment The inclusion of the domain ontology and its relation with the Thesauri of the DMS-1 standard facilitates the employment of this part of the model to support the manual annotation by providing lists of recommendations to the user based on the previous actions of other users. For some cases, the annotation tool only allows some concrete values for some information fields. The feeding of these dropdown menus is also done by employing the domain ontology.

5 Conclusions and future work

In this paper we have presented the semantic model implemented in the metadata model component of the RUSHES system. After explaining the motivation of the model, its implementation and the main benefits of that implementation have been described. As stated in [16], the component that relies on this model is able to provide all the semantic services required by the RUSHES system, which indexes, search and retrieve raw material.

We have also explained how the model is mainly driven by a industrial standard, presenting, to our knowledge, the first partial OWL implementation of a DMS-1 ontology.

Regarding the future work, the OWL DMS-1 ontology is not complete, since some parts of the standard that were out of the scope of RUSHES (e.g. management of the rights

from exploitation point of view) were not implemented. Due to this, the next main steps are to complete the DMS-1 ontology and make it public.

6 Acknowledgements

The RUSHES project (FP6-045189) is funded by the European Commission (FP6 program). Furthermore, we would like to thank the people from the Basque Television EiTB [2] for their kindly support.

References

[1] J. A. G. Avilés, B. León, K. Sanders, and J. Harrison. Journalists at digital television newsrooms in Britain and Spain: workflow and multi-skilling in a competitive environment. *Journalism Studies*, 5(1):87–100, February 2005.

[2] W. Bailer and P. Schallauer. The detailed audiovisual profile: Enabling interoperability between mpeg-7 based systems. In *Proceedings of 12th Multimedia Modellinig Conference*, pages 217–224, Beijing, 2006.

[3] W. Bailer, P. Schallauer, and H. Neuschmied. Mpeg-7 detailed audiovisual profile. Technical report, Joanneum Research, 2007.

[4] J. Bhogal, A. Macfarlane, and P. Smith. A review of ontology based query expansion. *Information Processing and Management: an International Journal*, 43(4):866–886, July 2007.

[5] D. Core. The dublin core metadata element set: An american national standard. Technical report, National Information Standards Organization, 2007.

[6] M. Cox, L. Tadic, and E. Mulder. *Descriptive Metadata for Television*. Focal Press, 2006.

[7] A. J. Faaborg. *A Goal-Oriented User Interface for Personalized Semantic Search*. PhD thesis, B.A. Information Science Cornell University, 2003.

[8] A. Ferrara, L. A. Ludovico, S. Montanelli, S. Castano, and G. Haus. A semantic web ontology for context-based classification and retrieval of music resources. *ACM Trans. Multimedia Comput. Commun. Appl.*, 2(3):177–198, t 06.

[9] B. C. Grau and B. Motik. Owl 1.1 web ontology language: Model-theoretic semantics. Technical report, Oxford University, 2008.

[10] M. R. Group. Digital media asset management & workflow management in the broadcast industry: Survey & analysis. Technical report, Research And Markets, 2004.

[11] W. Haas and H. Mayer. MPEG and its relevance for content-based multimedia retrieval. *J. UCS*, 7(6):530–547, 2001.

[12] IRT. Bmf. Technical report, IRT, 2007.

[13] L. Kerschberg and D. Weishar. Conceptual models and architectures for advanced information systems. *Applied Intelligence*, 13(2):149–164, 2000.

[14] H. Knublauch, R. W. Fergerson, N. F. Noy, and M. A. Musen. The protégé owl plugin: An open development environment for semantic web applications. In *International Semantic Web Conference*, pages 229–243, 2004.

[15] M. Krishna and D. Heppe. Dam and automated digital video workflows. *Journal of Digital Asset Management*, 2:163–167, 2006.

[16] G. Marcos, P. Krmer, A. Illarramendi, I. Garca, and J. Flrez. Semantic middleware to enhance multimedia retrieval in a broadcaster. In D. Luke, L. Hardman, A. Hauptmann, D. Paulus, and S. Staab, editors, *Semantic Multimedia, Third International Conference on Semantic and Digital Media Technologies, SAMT 2008 Koblenz, Germany, December 2008, Proceedings*, pages 74–88. Springer-Verlag Berlin Heidelberg, December 2008.

[17] S.-Y. Neo, J. Zhao, M.-Y. Kan, and T.-S. Chua. Video retrieval using high level features: Exploiting query matching and confidence-based weighting. In *Image and Video Retrieval*, pages 143–152. Springer Berlin / Heidelberg, 2006.

[18] A. G. Pérez and M. S. B. Galindo. *Reflections on language use in the academic context*, chapter Ontologies and Natural Language Processing, pages 121–161. Universidad Politécnica de Madrid, 2005.

[19] S. Pfeiffer and U. Srinivasan. Tv anytime as an application scenario for mpeg-7. In *MULTIMEDIA '00: Proceedings of the 2000 ACM workshops on Multimedia*, pages 89–92, New York, NY, USA, 2000. ACM.

[20] P. Schallauer, W. Bailer, and G. Thallinger. A description infrastructure for audiovisual media processing systems based on mpeg-7. *Journal of Universal Knowledge Management*, 1(1):26–35, 2006.

[21] O. Schreer, L. F. Ardeo, D. Sotiriou, A. Sadka, and E. Izquierdo. User requirements for multimedia indexing and retrieval of unedited audio-visual footage - rushes. In *Proc. of 9th Int. Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS)*, 2008.

[22] M. Shamsfard, A. Nematzadeh, and S. Motiee. Orank: An ontology based system for ranking documents. *International Journal of Computer Science*, 1:225–231, 2006.

[23] N. Simou, C. Saathoff, S. Dasiopoulou, E. Spyrou, N. Voisine, V. Tzouvaras, I. Kompatsiaris, Y. S. Avrithis, and S. Staab. An ontology infrastructure for multimedia reasoning. In L. Atzori, D. D. Giusto, R. Leonardi, and F. Pereira, editors, *VLBV*, volume 3893 of *Lecture Notes in Computer Science*, pages 51–60. Springer, 2005.

[24] P. J. Smith and Schirling. Metadata standards roundup. *IEEE Multimedia*, 13:84–88, April-June 2006.

[25] C. G. M. Snoek, B. Huurnink, L. Hollink, M. de Rijke, G. Schreiber, and M. Worring. Adding semantics to detectors for video retrieval. *IEEE Transactions on Multimedia*, 9(5):975–986, August 2007.

[26] T. Weithöner, T. Liebig, M. Luther, S. Böhm, F. von Henke, and O. Noppens. *The Semantic Web: Research and Applications*, chapter Real-World Reasoning with OWL, pages 296–310. Springer, 2007.

[27] J. Wiley and L. Sons, editors. *Introduction to MPEG-7, Multimedia Content Description Interface*. B. S. Manjunath, P. Salembier and T. Sikora (editors), Jun 2002.

[28] J. Wilkinson. Material exchange format (mxf) descriptive metadata scheme - 1 (standard, dynamic). Technical report, Pro-MPEG, 2003.

[2] EiTB Broadcaster website, http://www.eitb.com/