

Why Real-World Multimedia Assets Fail to Enter the Semantic Web

Tobias Bürger^{*}
Digital Enterprise Research Institute (DERI)
Technikerstrasse 21a
6020 Innsbruck, Austria
tobias.buerger@deri.at

Michael Hausenblas
Joanneum Research
Steyrergasse 17
8010 Graz, Austria
michael.hausenblas@joanneum.at

ABSTRACT

Making multimedia assets on the one hand first-class objects on the Semantic Web, while keeping them on the other hand conforming to existing multimedia standards is a non-trivial task. Most proprietary media asset formats are binary, optimized for streaming or storage. However, the semantics carried by the media assets are not accessible directly. In addition, multimedia description standards lack the expressiveness to gain a semantic understanding of the media assets. There exists an array of requirements both regarding media assets, and the Semantic Web already. Based on a critical review of these requirements we investigate how ontology languages fit into the picture. We finally analyse the usefulness of formal accounts to describe spatio-temporal aspects of multimedia assets in a practical context.

Categories and Subject Descriptors

H.5.1 [Information Systems]: Multimedia Information Systems; I7.4 [Document and Text Processing]: Electronic Publishing

General Terms

Multimedia Semantics, Semantic Web

Keywords

Multimedia Assets for the Semantic Web, Multimedia Models, Requirements Analysis

1. INTRODUCTION

Today a huge explosion of content can be experienced on the Web generated by, and for the home users [33]: An increasing number of people produce media assets (as photos, video clips, etc.), and share them on popular sites as Flickr¹, and YouTube².

*Tobias Bürger is also affiliated with Salzburg Research

¹<http://www.flickr.org>

²<http://www.youtube.com>

More recently, the popular attraction was guided away from image sharing to richer content sharing of videos. This can be seen by the launch of video portals like iFilm.com, Zidio.com or the dozen of other portals that appeared recently to compete with YouTube.

Unsurprisingly there is already a portal called VideoRonk³ trying to combine other portals by providing a MetaSearch interface, which is quite of an help as one does not want to search on ten or more different sites. However, what is missing is the link between the contents of all these sites, enabling distributed recommendations, cross-linking, etc.

Still, for example a cross-site search on the semantic level is close to impossible. The most obvious reason is due to a lack of metadata coming along with all the content. The power of providing metadata along with content on the Web can be seen at prospering mashups that not just combine APIs—provided by parties as Google⁴—but also trying to mashup things on a semantic level. This can be observed for example at Joost [40]. Having metadata about everything, as video content, blog posts, news feeds and the users of the system makes this new experience of watching TV through the Internet possible. To take this even one step further: Would every stream or video available on the Internet be described more detailed even content on the Internet could be matched with user profiles from applications like Joost and could be offered to watch.

As pointed out in [46, 36], high-quality metadata is essential for multimedia applications. Our recent work within initiatives [47] and research projects⁵ has shown, there is a need for going beyond current metadata standards to annotate media assets.

Current XML-based standards [24] are diverse, often proprietary and not ad hoc interoperable; cf. also [45]. In SALERO, for example, we are facing the problem to offer a semantic search facility over a diverse set of multimedia assets, e.g., image, videos, 3D objects or character animations. The same is true for the Austrian project GRISINO⁶ where we aim to realize a semantic search facility for cultural heritage collections.

³<http://www.videoronk.com>

⁴<http://code.google.com/apis/>

⁵such as in the EU project SALERO,
<http://www.salero.info>

⁶<http://www.grisino.at>

Automating the handling of metadata for these collections and automating linkage between parts of these collections is hard as the vocabularies to describe them are mostly diverse and do not offer facilities to attach formal descriptions.

A Motivating Scenario. Imagine a person that wants to watch the recent clips similar to the ones of his favourite experimental artist. Tons of clips are potentially distributed on the Web making a search for them time consuming and laborious. Thus a central facility to search for and negotiate content is needed. This facility should allow to formulate a search goal, including the characteristics, the subject matter, a maximum price, and the preferred encoding and file format of the clip. In a next step, all portal offerings will be scanned in order to retrieve and negotiate content that matches the users' intention. Note that also parts of a video may match his intention, which means that videos need to be fine granular and sufficiently well enough described.

In order for this scenario to work, the descriptions of (1) the goal formulation, (2) the description of the media content by all content owners and (3) the negotiation semantics have to be compatible. Three important focal points of these semantic descriptions are:

- *Expressivity for high level semantic descriptions of content* as typical users are not thinking in terms of colour histograms and spatial / temporal constructs. The characteristics of the media should be described detailed enough.
- The *need for rules*: To effectively identify the part of the content that matches the users' intention, rules are needed to map high level semantic concepts to spatial and temporal segments of the video (eg., because ratings and classifications could only apply to parts of the content, ie., a scene including crime is only suitable for adults)
- *Fine grain semantic descriptions* as of bandwidth, user effort, or cost reason to transfer the whole content is not possible. Thus parts of the content should be described detailed enough.

To reach out, we want to provide answers to the question: Why do we need rich semantic descriptions of media assets on the Web, and (why) is there a need to bundle these descriptions together with the multimedia assets? Simultaneously, we want to provide answers to the questions: How can descriptions be provided? Why are the metadata features of multimedia standards not enough?

Consequently, we elaborate on the answer to the question stated in the title of this paper “*Why Multimedia Assets Fail to Enter the Semantic Web?*” by first collecting the requirements for the description of multimedia assets (section 2), secondly by analysing the environment (section 3), and thirdly by collecting requirements for multimedia assets on the Semantic Web (section 4). In section 5 we analyse existing ontology languages for their usefulness regarding the requirements and conclude in section 6 with a discussion of the question stated in the title of this paper and give a brief outlook on the open issues.

2. REQUIREMENTS FOR THE DESCRIPTION OF MULTIMEDIA ASSETS

Requirements for multimedia content descriptions have been researched in a number of papers [17, 46, 36, 6] before and investigations of the combination of multimedia descriptions with features of the Semantic Web are yet numerous [27, 3, 42, 44, 2]. In the following, we give a summarisation of the proposed requirements and add two additional ones (Authoring & Consumption and Performance & Scalability).

Representational Issues. A basic prerequisite is the formal grounding and neutral representation of the format used to describe multimedia assets.

- *Neutral Representation*: The ideal multimedia metadata format has a platform and application independent representation, and is both human and machine processable;
- *Formal Grounding*: Knowledge about media assets must be represented in formal languages, as it must be interpretable by machines to allow for automation.

Extensibility & Reusability. It is requested that the format at hand is extensible, e.g., via an extension mechanism as found in MPEG-7. It should be possible to integrate or reference existing vocabularies [24].

Multimedia Characteristics and Linking. The format should reflect the characteristics of media assets, hence allow linking between data and annotations:

- *Description Structures*. The format should support description structures at various levels of detail, including a rich set of structural, cardinality, and multimedia data-typing constraints;
- *Granularity*. The language has to support the definition of the various spatial, temporal, and conceptual relationships between media assets in a commonly agreed-upon format;
- *Linking*. It has to facilitate a diverse set of linking mechanisms between the annotations and the data being described, including a way to segment temporal media.

Authoring & Consumption. A major drawback of existing metadata approaches is its lacking support for authors in creating annotations along with the lacking benefits of generated annotations.

- *Engineering support*. Appropriate tools are a prerequisite for uptake of new vocabularies. There is the need for at least authoring and consumption environments making use of the vocabularies to demonstrate their usefulness.

- *Deployment.* Multimedia Assets need to be exchangeable, and there must be ways to deploy descriptions along with the assets.

Performance & Scalability. The language should yield descriptions that can be stored, processed, exchanged and queried effectively and efficiently.

MPEG-7. MPEG-7 [35] is a powerful and flexible way to describe media assets at several levels of granularity; on the other hand MPEG-7 bears some intrinsic complexity and interoperability issues [4, 46, 36, 43]. Due to the fact that MPEG-7 standard is not grounded on formal semantics for the descriptions, variability in the syntactic representation of the descriptions may cause interoperability issues.

3. ENVIRONMENT ANALYSIS: THE SEMANTIC WEB

A good starting point for the analysis of our targeted hosting environment—the Semantic Web—is the Architecture of the World Wide Web [28], in which its three main building blocks are discussed: *identification*, *interaction*, and *data formats*. The Semantic Web, as an extension of the well-known Web roughly has the following characteristics:

- It is a *highly distributed system*. *Identification of resources* is based on URIs—for both data and services;
- There is no single, central “registry”, viz. *authorities are decentralised*; data and metadata are under control of a lot of distinct individuals (companies, standardisation bodies, private, etc.)
- Alike in the Web fundamental building blocks are *relations* between data, whereas the relations in the Semantic Web are *named*, may be of *any granularity* and allow the automatic interchange of data;
- *Contribuser*⁷ inhabit it; each participant may play different roles at once: consuming content and contributing via comments, links, etc.
- Finally, there exists a number of standards. Such as RDF allowing formal definitions of the intended meaning, SPARQL for querying, RDF(S), OWL or SKOS to classify content and OWL, WSML, or RIF for describing logical relationships.

Any multimedia metadata format that is after the successful application on the Semantic Web has to be in-line with the above listed characteristics. While some requirements, as formats (e.g. XML) are rather easy to meet, other can pose serious problems regarding the integration into the Semantic Web.

4. MULTIMEDIA ASSETS ON THE SEMANTIC WEB

Firstly, addressing the environmental requirements together with an efficient layering of the semantic descriptions on top

⁷a portmanteau word; contributor and user

of the existing metadata (sub-symbolic level - symbolic level - semantic level) is a necessary prerequisite for multimedia assets to enter the Semantic Web successfully. Secondly, from the requirements gathered in section 2 and the environmental analysis done in section 3 we deduce the following characteristics for multimedia assets on the Semantic Web:

Formality of Descriptions. Formal descriptions are the basic building blocks of the Semantic Web. To enable automatic handling like retrieval, and negotiation of multimedia assets formality of descriptions is a pre-requisite.

Three different (semantic) levels of multimedia metadata can be identified [17]: (1) At the *subsymbolic layer* covering the raw multimedia information typically binary formats are used which are optimized for storage or streaming and which mostly do not provide metadata. (2) The *symbolical layer* provides an additional structural layer for the binary essence stream. For this level standards like MPEG-7, Dublin Core or MPEG-21 can be used. The semantics of the information encoded with these standards are only specified within each standards framework. (3) Therefore the *semantic and logical layer* is needed to provide the semantics for the symbolical layer. This layer should be formally described using ontology languages as proposed in this paper.

Efficient layering and referencing of descriptions. It is necessary to support different levels of meaning attached to multimedia assets, i.e., meaning at the bit-level, *traditional metadata* and semantic (high-level) information. As there are already widely adopted standards available for the description of multimedia assets, the semantic layer must be efficiently put upon those traditional description layers and should not aim to replace it. Furthermore *semantic* descriptions from these traditional layers shall be re-used. As content, parts of content, and traditional and semantic descriptions may be distributed, efficient referencing mechanisms for multimedia content must be present.

Based on recent discussions⁸ we give a summarisation of possible approaches in the following. The multimedia asset is denoted with A , for the multimedia metadata (M_3) format, such as MPEG-7, we write M , the ontology (language) is written as O , and finally an external reference mechanism⁹ are labelled with R . The linking is depicted with \hookrightarrow :

- $M \hookrightarrow A$. the content is referenced from the M_3 format; the ontology layer has to deal with it, separately;
- $M \hookrightarrow O$. The M_3 format references the ontology layer;
- $O \hookrightarrow M$. The ontology layer references the M_3 format;
- $O \hookrightarrow A$. The ontology layer references the content directly;
- $O, M \hookrightarrow^R A$. The ontology layer and the M_3 format use a common reference mechanism to link to the content.

However, it has to be noted that there is no standardised way for the layering or the referencing, yet.

⁸<http://lists.w3.org/Archives/Public/public-xg-mmsem/2007Apr/0002.html>

⁹http://www.annodex.net/TR/URI_fragments.html

Interoperability among descriptions. Many formats used in various communities cause interoperability problems when dealing with multimedia content. To overcome this, an RDF based semantic layer should be added on top of these numerous formats to ease their semantic and syntactic integration. However, there are some open problems regarding the integration of existing annotation standards and semantic approaches [46, 36]: The stack of Semantic Web languages and technologies provided by the W3C is well suited to the formal, semantic descriptions of the terms in a multimedia document’s annotation. But, as also pointed out in [41], the Semantic Web based languages lack the structural advantages of the XML-based approaches. Additionally, there is a huge amount of work already done on multimedia document annotation within the framework of other standards. This is why a combination of the existing standards is the most promising path for multimedia document description in the near future.

Subjectivity and granularity of descriptions. Opinions and views on the content differ among users because of their personal background, culture or previous experiences. As many users are potential contributors to descriptions of assets, opinions may differ. Many of these opinions sometimes do not serve to a unique whole opinion. This is why it should be possible to separately attach these opinions to multimedia assets and keep them separate.

Trust and IPR issues. The Web consists of decentralized authorities and a huge number of contributers. As descriptions of content—especially in the new changing Web 2.0 environment—are subject to vandalism, there need to be ways to guarantee the validity of the descriptions and to secure descriptions that are just read-only for a user group. Popular portals like Flickr or YouTube show that there is no need to own content in order to annotate it. Furthermore copyright is critical when dealing with multimedia content.

Functional Descriptions. Sometimes the fact that metadata is created to support some specific function is forgotten when summarizing the requirements for a metadata schema. For the metadata creator it should be clear beforehand for what purpose the metadata will be used and what benefits he gains from it [34], ie., using this part of the metadata scheme enhances retrieval, raises social attention or helps you protect your assets.

This in turn also applies to the consumer of the metadata, functional descriptions of what type of information can be inferred from the attached metadata or what type of actions can be performed on the content are essential: this is especially true for information that is obfuscated prior to a possible negotiation phase of the content.

Engineering Support. The presence of metadata is a prerequisite to make multimedia assets accessible, and deployable on the Semantic Web, hence to enable their automated processing. From a developers perspective, there must be tools and standards enabling an integrated authoring, test-

ing, and deployment of multimedia assets along with their associated metadata. In the following the most important areas of engineering support are listed:

- *Edit & Visualise.* To aid the engineer in handling the annotations, editor tools, and IDEs¹⁰ are needed. These may include validator services¹¹, converter or mapper, and visualisation modules.
- *Libraries & Applications.* When developing applications, the availability of APIs is a core requirement. In special for Semantic Web applications, interface and mapping issues are of importance [19].
- *Deployment* Multimedia containers as HTML, SMIL, etc. require the metadata either being referenced from within the media assets, or being embedded into it. As the data model needs to be RDF—in contrast to existing, flat (tags, etc.) technologies—upcoming approaches as RDFa [1] need to be utilised thoroughly.

5. FORMAL DESCRIPTIONS OF MULTIMEDIA ASSETS

In this part ontology languages which are thought to be used for the advanced requirements which were identified in the sections before are introduced. In its core it comprises a comparison of two families of ontology languages against the requirements postulated in section 4.

The reader is invited to note that not all of the existing languages have the same expressiveness and not all have the same inferential capabilities. Further, the underlying knowledge representation paradigms can differ (eg., Description Logics, Logic Programming, etc.). Corcho and Gomez-Perez [20] present a framework that allows for analysing and comparing the expressiveness and reasoning capabilities of ontology languages, which can be used in the decision process. The process of choosing and selecting the appropriate ontology language includes questions about e.g. the expressiveness, inference mechanisms, translators or exchange formats offered for an ontology language. We are going to take these questions into consideration and simultaneously verify if the languages meet the requirements discussed in section 4.

5.1 Ontology Languages

A number of logical languages have been used for the description of different kinds of knowledge (i.e. ontologies and rules) on the Semantic Web: First Order Logic, Description Logics, Logic Programming and Frame-based Logics. Each of which allow the description of different statements and each imply different complexity results for certain reasoning tasks with these languages.

In this section we want to introduce two of the most promising ontology language families, ie., the OWL- and the WSMIL-family of languages. The OWL family of languages is a standardisation effort of the W3C and the WSMIL family of languages is an effort of the WSMO working group, whereas WSMIL is a formal language for the description of ontologies and Semantic Web Services. Other ontology languages like

¹⁰as for example <http://www.topbraidcomposer.com/>

¹¹<http://phoebus.cs.man.ac.uk:9999/OWL/Validator>

F-Logic [30], OIL [7] or DAML+OIL¹² were not taken into consideration because their lack of support for recent W3C recommendations like RDF.

5.1.1 Web Ontology Language (OWL) Family

The Web Ontology Language (OWL) family was designed in a W3C standardisation process because of the need for an ontology language that can be used to formally describe the meaning of terminology used in Web documents, thus, making it easier for machines to automatically process and integrate information available on the Web. This language should be layered on top of XML and RDF (W3C's Resource Description Framework¹³) in order to build on XML's ability to define customized tagging schemes and RDF's approach to representing data.

Currently OWL 1.1¹⁴ is under development; it extends OWL DL in several ways: the underlying DL now is *SROIQ*, which provides increased expressive power with respect to properties and cardinality restrictions. Further, OWL 1.1 has user-defined datatypes and restrictions involving datatype predicates, and a weak form of meta-modelling known as punning.

The usage of rules in combination with DL has been investigated for some time [14, 21]—in the Semantic Web stack, it is expected that a rule language will complement the ontology layer.

5.1.2 The WSML family of languages

The activities of the WSMO Working group¹⁵ have yielded proposals of new ontology languages, namely WSML (WSML-Core, WSML-DL, WSML-Flight, WSML-Rule, WSML-Full), OWL- ("OWL minus") [8] and OWL Flight [10]. In [16] unique key features of WSML in comparison of other language proposals are presented. Compared to OWL key features include (1) WSML offers one syntactic framework for a set of layered languages, and (2) it separates between conceptual and logical modelling. An overview of the different variants of the WSML framework can be found in [32]. One has to note that WSML-Flight incorporates a rule language while still allowing efficient decidable reasoning and WSML-Rule allows unsafe rules. The relation of WSML to OWL is discussed in [9].

5.2 Rules

Due to well-known limitation of the expressive power of the Description Logics language family [25, 26], the need for a richer set of descriptions w.r.t. properties emerges. As rule systems are widely deployed, the harmonisation efforts have not been successful so far. A relatively new W3C initiative, the Rule Interchange Format Working Group, is now after defining a core rule language for exchanging rules. This Rule Interchange Format Core¹⁶ (RIF Core) language aims at achieving maximum interoperability while preserving rule

¹²DAML+OIL Reference Description, see:
<http://www.daml.org/2001/03/reference>

¹³<http://www.w3.org/TR/rdfprimer/>

¹⁴http://owl1_1.cs.manchester.ac.uk/owl_specification.html

¹⁵<http://www.wsma.org>

¹⁶<http://www.w3.org/TR/rif-core/>

semantics; from a theoretical perspective, RIF Core corresponds to the language of definite Horn rules. As standardisation is still in its infancy, we will not go further into detail regarding rules, but one has to note that the careful integration of ontology languages is an issue to be addressed; for example the usage of DL concepts in a rule has to be well-defined.

5.3 Comparing Formal Descriptions Regarding the Requirements

In the following a high-level comparison of formal description paradigms for multimedia assets is performed. We chose OWL+RIF on the one side, and WSML/OWL-Flight on the other to achieve a somehow realistic scenario; the result can be found in Table 1¹⁷: The table indicates for which requirement an ontology language (resp. OWL / WSML) can be utilised to overcome the identified shortcomings of traditional approaches and thus fulfill the requirements stated in 4.

Requirement	OWL 1.1 + RIF	WSML-/OWL-Flight
Formal Description	++	++
Layering of Descriptions	+	+
Interoperability	++	+
Granularity	-	-
Trust & IPR issues	-	-
Functional Descriptions	-	*
Engineering Support	++	+
Datatype Support	+	++

Table 1: Comparison of Formal Descriptions for Media Assets.

In the following, we elaborate in detail on each of the items in Table 1, and argue therefore our findings regarding the comparison of OWL 1.1 + RIF vs. WSML/OWL-Flight.

5.3.1 Formal Description

Both OWL and WSML provide a framework for the formal (machine-processable) description of ontologies. An ontology in WSML consists of the elements concept, relation, instance, relationInstance and axiom. The primary elements of an OWL ontology concern classes and their instances, properties, and relationships between these instances. The formality of the descriptions is based on logics that allow machines to reason on the information. Whereas OWL is based on Description Logics, the WSML family members are based on different logic languages (ie. Description Logics, Logic Programming or First Order Logic).

Despite the fact, that OWL is more widely adopted and used we believe that WSML with its layered framework is conceptually superior to OWL. A major difference between ontology modeling in WSML and ontology modeling in OWL

¹⁷++ ... good support, + ... available, - ... not supported,
* ... supported because of WSML's constructs for the description of Semantic Web Services

is that WSMC separates conceptual modelling for the non-expert users, and logical modeling for the expert user as it—unlike OWL—uses an epistemology, which abstracts from the underlying logical language making the *surface syntax* nicer. Even if an application later requires OWL, one is able to use WSMC tools to convert ontologies that reside in popular logic/language fragments automatically into equivalent OWL ontologies. Furthermore the WSMC family framework enables one to choose exactly the language with the needed expressiveness to be used, and later allows an easy switch to another family member as a consequence of its common grounding. WSMC Rule and WSMC Flight also include rule-support. Thus, unlike with OWL, no additional rule language is needed.

5.3.2 Layering of Descriptions

An array of existing multimedia metadata (M3) formats have been used for years in diverse application areas. However, when one aims at using these formats (as MPEG-7, ID3, etc.) in the context of the Semantic Web, the options are limited. Hence, to enable an efficient layering of RDF-based vocabularies on top of existing multimedia metadata, one may use hybrid techniques.

As a result of our works in the media semantics area, we recently proposed the RDFa-deployed Multimedia Metadata (**ramm.x**) specification [22]. **ramm.x** is a light-weight framework allowing existing multimedia metadata to hook into the Semantic Web using RDFa [1]. Ontologies based on WSMC and OWL are typically used in **ramm.x** to formalise a M3 format; this is especially important due to their interoperability features (see 5.3.3).

A different but as well Web compatible approach is described in [31]. There, the authors propose the concept of semantic documents; semantic documents include any information regarding the document and its relationships to other documents. The concept is realised by including XMP descriptions in PDF documents which can be rendered in any browser with available plugins. XMP is a format for embedding metadata in documents using RDF.

5.3.3 Interoperability

To adhere to the architecture of the WWW, OWL uses (1) URIs for naming and (2) RDF to provide extensible descriptions. (3) OWL builds on RDF and RDF Schema and adds additional vocabulary for describing properties and classes. (4) The datatype support for OWL is grounded on XML Schema.

WSMC has a number of features which allow to integrate it seamlessly in the Web: (1) WSMC uses IRIs¹⁸ [15] for the identification of resources. (2) WSMC adopts the namespace mechanism of XML, and WSMC and XML Schema datatypes are compatible. (3), WSMC has an XML- and RDF based syntax for exchange over the Web.

To reach compatibility between WSMC and OWL, WSMC has a set of defined translators between OWL and WSMC [11, 12].

¹⁸IRIs are the successors of URIs

5.3.4 Granularity

As stated above, when referring to granularity, we understand the support of the definition of various spatial, temporal, and conceptual relationships regarding annotations. In this sense, OWL and WSMC meet the minimal requirements, but do not explicitly address this issue. Depending on the granularity, obviously scalability and performance issues come along. In this respect, again, OWL and WSMC can be perceived comparable.

5.3.5 Trust and IPR

In an interdependent, interconnected environment as the Semantic Web, two important aspects immediately arise: data provenance and trust [5]. Requirements regarding trust issues gathered from [37, 18] contain costs and benefits w.r.t. implementation, technology-driven vs. social networking, etc.

Both WSMC and OWL do not have explicit provisions for handling trust and IPR issues, respectively.

5.3.6 Functional Descriptions

OWL and the WSMC’s part for the description of ontologies do not have support for such kind of descriptions. However, WSMC is a language for the specification of ontologies and different aspects of Web services. As such it not only provides means for modeling and description of ontologies but also functional (service) descriptions, i.e. the description of a service capability by means of precondition, assumptions, postconditions and effects [29].

5.3.7 Engineering Support

Tool Support for WSMC and especially OWL is constantly growing. However, the amount of tools available for OWL [48] and WSMC [13] can drastically not be compared. As OWL is a W3C Recommendation, the support for it is huge.

5.3.8 Data Type Support

The reader is invited to note that both OWL and WSMC ground their datatype support on XML Schema. In WSMC, XML Schema primitive datatypes, simple types and XML Schema derived datatypes are supported [39]; OWL adopts the RDF(S) specification of datatypes [38], though some XML Schema built-ins are problematic.

6. CONCLUSIONS & OUTLOOK

The first question we kept open is "What are real-world multimedia assets"? Real-world multimedia assets are multimedia objects which can be currently found embedded in HTML pages on the Web, as images, videos, etc. We see three main reasons why media assets fail to enter the Semantic Web:

1. There is a lack of the critical mass of annotated content which is mainly due to the large scale automation of (semantic) visual analysis has not gone that far. This is why the user is the central person in the process in order to provide manual annotations. Motivating user to attach complex annotations to content is not easy to achieve.

2. Current traditional and Web 2.0 based approaches to multimedia annotation are not useful to achieve the goals of the Semantic Web: The most important aspects that the Semantic Web intends to solve are (i) Annotation, (ie., how to associate metadata to a resource), (ii) Information Integration (ie., how to integrate information about resources), and (iii) Inference (ie., reasoning over known facts to unleash hidden facts).
- Existing multimedia metadata standards as MPEG-7 can be used to annotate but keep a certain amount of ambiguity amongst these annotations. As it is a standard it allows easy integration based on it (a requirement for that is that everyone adheres to this standard!) but inference is not possible with the information attachable to a MPEG-7 file. The problem with tagging is manifold; there are open issues, such as consistency among tags, reconciliation of tags, and how to associate tags with parts of the tagged content. This huge amount of uncertainty will not allow reliable information integration, nor allow to reason on it.
3. As we argued in this paper, more requirements have to be fulfilled, which can not be solely solved by traditional or Web 2.0 based approaches and which make more formalized descriptions of content necessary. However, before not being able to attach these directly to the media being described, multimedia assets will not be able to enter the Semantic Web.

Regarding deployment of M3 format on the Semantic Web, we recently proposed to use ramm.x in the Cultural Heritage domain [23].

Acknowledgements

The research leading to this paper was partially supported by the European Commission under contract FP6-027026, "Knowledge Space of semantic inference for automatic annotation and retrieval of multimedia content - K-Space" and SALERO (contract number FP6-027122).

7. REFERENCES

- [1] B. Adida and M. Birbek. RDFa Primer 1.0 - Embedding RDF in XHTML. W3C Working Draft, W3C RDF in XHTML Taskforce, 2007.
- [2] R. Arndt, R. Troncy, S. Staab, L. Hardman, and M. Vacura. COMM: Designing a Well-Founded Multimedia Ontology for the Web. In *Proceedings of the 6th International Semantic Web Conference (ISWC'2007), Busan, Korea, November 11-15, 2007, (forthcoming)*, 2007.
- [3] T. Athanasiadis, V. Tzouvaras, K. Petridis, F. Precioso, Y. Avrithis, and Y. Kompatsiaris. Using a Multimedia Ontology Infrastructure for Semantic Annotation of Multimedia Content. In *Proc. of 5th International Workshop on Knowledge Markup and Semantic Annotation (SemAnnot '05), Galway, Ireland, November 2005*, 2005.
- [4] W. Bailer, P. Schallauer, M. Hausenblas, and G. Thallinger. MPEG-7 Based Description Infrastructure for an Audiovisual Content Analysis and Retrieval System. In *Proceedings of SPIE - Storage and Retrieval Methods and Applications for Multimedia*, pages 284–295, San Jose, California, USA, 2005.
- [5] C. Bizer and R. Oldakowski. Using context- and content-based trust policies on the Semantic Web. In *Proceedings of the 132th international World Wide Web conference on Alternate track papers & posters*, pages 228–229. ACM Press, 2004.
- [6] T. Bürger and R. Westenthaler. Mind the gap - requirements for the combination of content and knowledge. In *Poster Proceedings of the SAMT 2006 Conference, Athens, Greece*, 2006.
- [7] F. D., van Harmelen F., H. I., M. D., and P.-S. P. Oil: An ontology infrastructure for the semantic web. *IEEE Intelligent Systems and their applications*, 16(2):38–44, 2001.
- [8] J. de Bruijn and A. P. (eds.). OWL⁻. WSM_L Deliverable D20.1v0.2 WSM_L Working Draft 05-15-2005, <http://www.wsmo.org/TR/d20/d20.1/v0.2/>, 2005.
- [9] J. de Bruijn, H. Lausen, A. Polleres, and D. Fensel. The Web Service Modeling Language WSM_L: An Overview. In *ESWC*, pages 590–604, 2006.
- [10] J. de Bruijn (ed.). OWL Flight. D20.3v0.1 OWL Flight WSM_L Working Draft 23-08-2004, <http://www.wsmo.org/2004/d20/d20.3/v0.1/>, 2004.
- [11] DERI. OWL - WSM_L Translator v1.0. <http://tools.deri.org/wsml/owl2wsml-translator/v0.1/>, 2007.
- [12] DERI. WSM_L - OWL Translator v1.0. <http://tools.deri.org/wsml/wsml2owl-translator/v0.1/>, 2007.
- [13] DERI. WSM_L Tools. <http://tools.deri.org/wsml/>, 2007.
- [14] F. M. Donini, M. Lenzerini, D. Nardi, and A. Schaerf. AL-log: Integrating Datalog and Description Logics. *Journal of Intelligent Information Systems*, 10(3):227–252, 1998.
- [15] M. Duerst and M. Suignard. Internationalized Resource Identifiers (IRIs). IETF RFC 3987, 2005. <http://www.ietf.org/rfc/rfc3987.txt>.
- [16] D. Fensel, H. Lausen, A. Polleres, J. de Bruijn, M. Stollberg, D. Roman, and J. Domingue. *Enabling Semantic Web Services: The Web Service Modeling Ontology*. Springer, 11 2006.
- [17] J. Geurts, J. van Ossenbruggen, and L. Hardman. Requirements for practical multimedia annotation. In *Proceedings of the Workshop on Multimedia and the Semantic Web, May 2005, Heraklion, Crete*, pages 4–1, 2005.
- [18] J. Golbeck, B. Parsia, and J. Hendler. Trust Networks on the Semantic Web. In *Proceedings of Cooperative Intelligent Agents 2003*, 2003.
- [19] N. M. Goldman. Ontology-Oriented Programming: Static Typing for the Inconsistent Programmer. In *Proceedings of the Second International Semantic Web Conference - ISWC 2003*, pages 850–865, 2003.
- [20] Gomez-Perez, Fernandez-Lopez, and Corcho-Garcia. *Ontological Engineering*. Springer, Berlin, 2004.
- [21] B. Grosop, I. Horrocks, R. Volz, and S. Decker. Description Logic Programs: Combining Logic Programs with Description Logics. In *12th International World Wide Web Conference*

- (*WWW'03*), Budapest, Hungary, 2003.
- [22] M. Hausenblas, W. Bailer, and T. Bürger. Deploying Multimedia Metadata on the Semantic Web - RDFa-deployed Multimedia Metadata (*ramm.x*). Specification, *ramm.x* Working Group, 2007.
- [23] M. Hausenblas, W. Bailer, and H. Mayer. Deploying Multimedia Metadata in Cultural Heritage on the Semantic Web. In *First International Workshop on Cultural Heritage on the Semantic Web, collocated with the 6th International Semantic Web Conference (ISWC'07)*, Busan, South Korea, 2007.
- [24] M. Hausenblas, S. Boll, T. Bürger, O. Celma, C. Halaschek-Wiener, E. Mannens, and R. Troncy. Multimedia Vocabularies on the Semantic Web. W3C Incubator Group Report, W3C Multimedia Semantics Incubator Group, 2007.
- [25] I. Horrocks, P. F. Patel-Schneider, S. Bechhofer, and D. Tsarkov. OWL rules: A proposal and prototype implementation. *Journal of Web Semantics*, 3(1):23–40, 2005.
- [26] I. Horrocks, P. F. Patel-Schneider, and F. van Harmelen. From *SHIQ* and RDF to OWL: The making of a web ontology language. *Journal of Web Semantics*, 1(1):7–26, 2003.
- [27] J. Hunter. Adding Multimedia to the Semantic Web - Building an MPEG-7 Ontology. In *First International Semantic Web Working Symposium (SWWS'01)*, Stanford, California, USA, 2001.
- [28] I. Jacobs and N. Walsh. Architecture of the World Wide Web, Volume One.
<http://www.w3.org/TR/webarc/>, 2004.
- [29] U. Keller, H. Lausen, and M. Stollberg. On the Semantics of Functional Descriptions of Web Services. In *The Semantic Web: Research and Applications (Proceedings of ESWC 2006)*, pages 605–619, 2006.
- [30] M. Kifer and G. Lausen. F-logic: a higher-order language for reasoning about objects, inheritance, and schema. In B. L. J. Clifford and D. Maier, editors, *Proceedings of the 1989 ACM SIGMOD international Conference on Management of Data*, pages 134–146, New York, NY, 1989. ACM Press.
- [31] H. Kim, H. Kim, J. H. Choi, and S. Decker. Translating Documents into Semantic Documents using Semantic Web and Web 2.0. In *Proceedings of the 1st Semantic Authoring and Annotation Workshop (SAAW2006)*, 2006.
- [32] H. Lausen, J. de Bruijn, A. Polleres, and D. Fensel. WSML - a Language Framework for Semantic Web Services. In *Proceedings of the W3C Workshop on Rule Languages for Interoperability*, 2005.
- [33] J. Markoff. Web content by and for the masses. New York Times Online, June 2005.
- [34] A. Morgan and M. Naaman. Why we tag: motivations for annotation in mobile and online media. In *CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 971–980, New York, NY, USA, 2007. ACM Press.
- [35] MPEG-7. Multimedia Content Description Interface. Standard No. ISO/IEC n°15938, 2001.
- [36] F. Nack, J. van Ossenbruggen, and L. Hardman. That Obscure Object of Desire: Multimedia Metadata on the Web (Part II). *IEEE Multimedia*, 12(1), 2005.
- [37] K. O'Hara, H. Alani, Y. Kalfoglou, and N. Shadbolt. Trust Strategies for the Semantic Web. In *ISWC Workshop on Trust, Security, and Reputation on the Semantic Web*, 2004.
- [38] J. Z. Pan. *Description Logics: Reasoning Support for the Semantic Web*. PhD thesis, School of Computer Science, The University of Manchester, 2004.
- [39] D. Roman, H. Lausen, and U. Keller. Web Service Modeling Ontology (WSMO), WSMO Deliverable D2v1.0., WSMO Working Draft 20 September 2004, September 2004.
- [40] L. Simons. RDF at the Venice Project.
<http://www.leosimons.com/2006/rdf-at-the-venice-project.html>, 2006. Blog Post.
- [41] G. Stamou, J. van Ossenbruggen, J. Z. Pan, and G. Schreiber. Multimedia annotations on the semantic web. *IEEE MultiMedia*, 13(1):86–90, 2006.
- [42] R. Troncy. Integrating Structure and Semantics into Audio-visual Documents. In *Proceedings of the 2nd International Semantic Web Conference (ISWC'03)*, volume LNCS 2870, pages 566–581, 2003.
- [43] R. Troncy, W. Bailer, M. Hausenblas, P. Hofmair, and R. Schlatte. Enabling Multimedia Metadata Interoperability by Defining Formal Semantics of MPEG-7 Profiles. In *1st International Conference on Semantics And digital Media Technology (SAMT'06)*, pages 41–55, Athens, Greece, 2006.
- [44] C. Tsinaraki, P. Polydoros, and S. Christodoulakis. Integration of OWL ontologies in MPEG-7 and TVAnytime compliant Semantic Indexing. In *Proceedings of the 16th International Conference on Advanced Information Systems Engineering (CAiSE)*, 2004.
- [45] V. Tzouvaras (ed.). Multimedia Annotation Interoperability Framework; MMSEM XG Report.
http://www.w3.org/2005/Incubator/mmsem/wiki/Semantic_Interoperability, 2007.
- [46] J. van Ossenbruggen, F. Nack, and L. Hardman. That Obscure Object of Desire: Multimedia Metadata on the Web (Part I). *IEEE Multimedia*, 11(4), 2004.
- [47] W3C. Multimedia Semantics Incubator Group.
<http://www.w3.org/2005/Incubator/mmsem/>, 2007.
- [48] W3C. Semantic Web Development Tools.
<http://esw.w3.org/topic/SemanticWebTools>, 2007.