

Towards Increased Reuse: Exploiting Social and Content Related Features of Multimedia Content on the Semantic Web

Tobias Bürger

Semantic Technology Institute (STI),
University of Innsbruck, Austria,
tobias.buerger@sti2.at

Abstract. While the amount of multimedia content on the Web is continuously growing, reuse of multimedia content remains low and automated processing of content remains hard. An increased reuse of content however would result in a greater consistency, quality and lowered cost of the production of new content. The retrieval of multimedia content on the Web is a continuous challenge which is due to the lack of (formal) descriptions and generic multimedia analysis algorithms. We present a model and a set of ontologies to mark up multimedia content embedded in web pages which can be used to deploy its descriptions on the Semantic Web and which in turn can be used to reason about the contents of multimedia resources. This model is part of a method to raise the reusability potential of content which in turn is expected to lower production costs of new content.

1 Introduction

In 2001, Tim Berners-Lee et al. introduced their vision of an augmented Web in which information is meaningful for machines as well as for humans. Since the introduction of this vision, the Web more and more turned into a multimedia environment and became a place to share great amounts of professionally and user generated content.

One of the challenges brought forward by the Semantic Web is the need to enrich existing digital content in such a way that machines can determine what the content is about, how it can be used, and whether one needs to pay for it or not. This part of the vision, richly annotated and formally described content which supports its automated negotiation, is sometimes referred to as *Intelligent Content*. Not only the ever growing amount of digital content raised interest in this topic in the recent years, but also the lack of appropriate multimedia description standards for the explication of features of content [22, 16]. Relevant features include not only the semantic content of images but also structural, legal or behavioral issues as especially multimedia content has many characteristics that for different usage scenarios need to be described.

One question that we especially intend to answer is how multimedia content which is published on the Web can be described to efficiently be reused, republished or reformatted for different purposes or different target media. And more concrete: How can social and content-related features and descriptions of content on the Web be exploited to increase the reusability potential of content.

The contribution of this paper is a conceptual model to describe and represent multimedia content including its context on the Web. The model is especially designed to raise the reusability potential of content published in typical web pages (e.g. embedded in news stories), social media or professional image licensing sites. The model is supported by a set of ontologies to mark up multimedia resources inline of HTML pages using RDFa¹.

2 Motivating Scenario: Towards Web-scale Reusability of Content

Supporting the reuse of content can provide significant improvements in the way how content is created and used, including increased quality and consistency, long-term reduced time and costs for development, maintenance, or adaptation to changing needs [18]. The amount of content available on the Web grows every day and the amount of professionally produced content in local or commercial databases also stays on a high level all of which could potentially be reused. The wish for reuse of content comes inline with the need for automation of associated tasks like search & retrieval, selection, or adaptation of content. However, automated handling is mostly hindered by the fact that users search for content based on the aboutness of the contained information which is – if at all – represented by tags attached to the content on the Web. Still, high-level features which are of high importance for retrieval of content [13] are not automatically derivable by most analysis algorithms which is due to the Semantic Gap [19], which commonly refers to the large gulf between automatically extractable low-level features and high-level semantics which are typically derived based on the background of a human being. Furthermore licenses and conditions of use are mostly encoded in web pages using natural language which is understandable by humans but not by machines. With the aim to automate handling which includes selection of the right pieces of content, this fact demands for richer semantic descriptions of content. Having richer semantic descriptions in turn implies improvements in reuse and automation.

Currently content on the Web is published based on different metadata standards and with different intention in mind. End users either publish images for non-commercial aspects, i.e. for others to watch or to gain reputation, or they publish it out of a commercial reason, either because they want to sell it or to grant access in order to gain revenues.

Current Web-based content reuse is difficult because amongst others

- Presentations are mostly available as PDF files with some tags attached to them. However fine granular descriptions of images, which were used in some slides and which would be candidates for reuse, are mostly missing.
- The same is true for videos: Tags are provided which are often not covering the semantics of particular scenes but only of the whole video.
- Cross-site searches in commercial image libraries are often hindered by the fact that images are described differently across sites.
- A huge amount of sites are using images for illustrative purposes which are not explicitly described. Sometimes textual descriptions are provided with the images but these are not explicitly assigned to the image which again makes retrieval hard.

¹ <http://www.w3.org/TR/xhtml1-rdfa-primer/>

We intend to overcome these difficulties using a **two-fold strategy** in order to raise the reusability rate of content: (1) Raise the findability rate of content by unlocking the reusability potential of content published on social media, non-commercial or commercial sites by formally describing their content related- and social features and (2) Increase the ad-hoc adaptability of content by allowing to select and describe parts of it instead of only the singular datastream / resource.

3 Requirements for Intelligent Content with respect to Reusability

Content is an *individual securable and targeted reproduction of implicit information done by humans*. Important aspects of content, which are important with respect to reuse, include:

- the *contextual aspect* (ie. what the information or the content is about),
- the *technical aspect* (ie. how is it technically represented,)
- the *economic aspect* (ie. what is the value of the content), and
- the *legal aspect* (ie. what are the rights to use the content)

Intelligent Content Objects (ICOs) – as we understand them – are inspired by the vision of smart content objects [1], which define a package structure including the content, knowledge about its properties and several interfaces to interact with the smart content object. This is similar to packaging standards like MPEG-21 [4] or OAI Information Packages [7] with the difference that semantic technologies are explicitly used to provide machine understandable descriptions.

Requirements for multimedia content descriptions have been researched before and investigations of the combination of multimedia descriptions with features from the Semantic Web are yet numerous which we summarize in [5]. We want to highlight this issue again with respect to reusability of content which we believe deserves special attention.

We identified 3 aspects that need to be fulfilled in order to increase the reusability potential of content:

1. **Findability:** Reuse is often hindered by the fact that people are not aware of content to be re-used because it can not be found. This is especially true for multimedia content. Thus there is a need for a metadata model especially supporting findability and reusability of content. The model has to support descriptive information but also needs to support linking and referencing of secondary information and to acknowledge the fact that on the Web2.0 different users may provide metadata by tagging, rating, or referencing.
2. **Adaptability:** Resources published on the Web are mostly atomic. They are mostly available in a single file even if the file has been assembled out of different datastreams. This makes content difficult to repurpose. The existence of a reusability-friendly format that makes structure explicit could however enable the reuse of components as well. Thus there is the need for a model that allows to access sub-components and which enables to structure content and to identify and select its sub-parts.

3. **Cross-Community Interoperability:** Query mechanisms for content must reflect habits of people from different communities. People from the E-Learning domain are used to think in terms of learning objects and content fragments while people from the archival domain communicate in terms of Information Packages or Information Objects, etc. This demands for a basic compatibility with existing standards.

4 A Conceptual Model for Intelligent Content for the Semantic Web

In this section we present a reference model for intelligent content which takes the characteristics of the Semantic Web [5] and the characteristics of the Web 2.0 as being a paradigm for rich social interaction on the Web into account.

4.1 A Data Model for Intelligent Content

Firstly there is a need for an abstract model that offers a set of well-defined concepts and vocabularies to sketch the problem of how content can be described and how it supports *adaptability* and *findability* while remaining compatible with existing data models from the multimedia, E-Learning and archival domain to support *interoperability*.

The general aim is to lay a graph over published contents on the Web, associating descriptions in a Web page to it and to bind metadata to content and descriptive information. This approach is similar to the way how digital assets are organized in the information domain. Here digital assets aggregate multiple-streams of relevant data, descriptive metadata and secondary data into one compound object which is then managed by a single entity (cf. section 6).

In [3] Boll et al. compared multimedia document models according to advanced requirements for reusability, adaptability and usability from a technical perspective. Important aspects regarding reusability are: (1) Granularity of media elements, fragments and documents (2) Kind of reuse, i.e. structural or identical, and (3) Identification and selection. Based on an assessment of different models according to these characteristics we selected the MPEG-21 Digital Item Declaration (DID) Abstract Model [4] as a data model that fulfills the basic characteristics of an adaptable data model, i.e. granular description of fragments, media elements (resources), grouped resources (i.e. components) and identification and selection of (parts of) resources. The basic parts of an MPEG-21 Digital Item which are interesting from this aspect are depicted in Figure 1: These include most notably containers in which identifiable digital assets are included (items) and which contain (multimedia) resources. Fragments of these resources can be selected and both resources and their fragments can be described via descriptors or annotated via annotations. Our proposed data model is realized by an ontology which covers the MPEG-21 DID Abstract Model and which amongst others makes the semantic types of relations between media elements, components or fragments and their descriptors explicit. The ontology is briefly described in section 4.3. The MPEG-21 DID Abstract Model is amongst other compatible with the OAI Abstract Information Model as shown in [2] and our first investigations also indicate that it is compatible to existing learning content models.

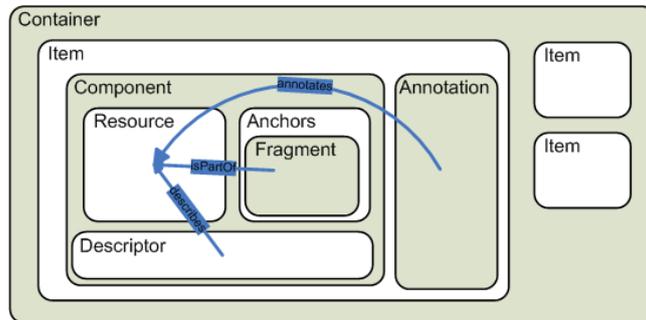


Fig. 1. MPEG-21: Main elements within the Digital Item Declaration Model

4.2 A Metadata Model to Increase the Findability Rate of Content

Multimedia retrieval is the discipline of applying information retrieval techniques to non-text-based content. The critical point for these techniques is, that users search for non-text based content based on the aboutness of the contained information [13] which is – if at all – represented by tags attached to the content or by information derived from the deployment context of the content which – when used for search – is blurred in the retrieved result sets. This is why reliable metadata is often essential to enable retrieval of multimedia content. It is commonly acknowledged that a metadata model which increases the findability rate of content also increases the reusability potential of content [18]. We follow a **two-fold strategy to increase the findability rate of content**: (1) First descriptions provided along with the content on web sites are explicitly related with the content to provide hints for search engines where to find information that can be used for indexing. This is supported by the data model introduced in section 4.1. (2) Secondly, content will be accompanied by metadata sets following a metadata model which captures relevant social and content related features and which is outlined in the subsequent section.

Social Aspects: The Role of Different Users in Content and Metadata Production

We analysed the life cycle of both content and metadata and the roles of distinct user groups in order to determine components that our metadata model has to provide. Here, we especially took social aspects of content into account which is an important indicator for reuse. The content lifecycle consists of the following different dimensions (cf. [14]):

1. **The User Dimension:** Content and metadata is produced, altered and consumed by different users playing different roles: Production related users who create, process, resell, or publish content and end users who mainly consume but also increasingly produce content.
2. **The Content Dimension:** During its lifetime content is transferred between different stages. According to the canonical processes of media production [10] it is premediated, created (processed), annotated, packaged, organized and distributed.
3. **The Metadata Dimension:** Metadata is potentially being added by different users in every step of the content lifecycle. In [9] four different metadata creation roles

are introduced: The content creator who directly provides metadata, professional metadata creators who get paid for annotating content, technical metadata creators who just add basic technical metadata, and community enthusiasts which are very prominent in the Web 2.0 and tag content.

From this observation we are able to derive that metadata about content is not static and should be changeable during lifetime. Furthermore metadata is potentially being provided by different parties.

Unblurring Content Descriptions: Supporting Multiple Metadata Sets In the literature the distinction is made between authoritative and non-authoritative metadata. Authoritative metadata is contributed by the author (creator) of the content and reflects persistent information about the content. Non-authoritative metadata is provided by the consumer or a third party and provides contextual and changing aspects [17]. Non-authoritative metadata is especially useful for recommendations based on collaborative filtering techniques and thus is critical in the effective discovery and reuse of content. To reflect different opinions and interpretations of content, metadata provided by different parties should be connected to its originator which is an important indication of its quality and trustworthiness and thus should be kept separate. Our model therefore explicitly supports *one authoritative metadata and multiple non-authoritative sets to be attached to different parts of the data model*. While authoritative metadata is explicitly added, non-authoritative metadata may explicitly be provided (through annotations, reviews, ratings, etc.) or implicitly be generated (through harvesting or usage analysis).

Types of Metadata As previously said, metadata is critical for the discovery of non-text based content on the Web. Metadata standards or vocabularies for multimedia are yet numerous as we summarize in [12].

Our investigations of standards and types of metadata focused on a core set which reflects the properties of content in its lifecycle. This set mostly covers the core facets which are used across a variety of domains and which we believe are important to support findability with respect to reuse:

1. **Bibliographic metadata** is traditionally concerned and related to the authorship of content and includes basic fields like identification, naming, publication or categorization.
2. **Technical metadata** typically describes physical properties of content, like format, bit-rate and what is mostly called low-level features of content.
3. **Classification metadata** might include keywords or tags but also domain specific classification information
4. **Evaluative metadata** includes ratings and qualitative assessments of the content. A collaborative evaluation model based on evaluative metadata could provide invaluable information regarding reuse of content. This could include dimensions like usefulness, presentation aesthetics, or design.
5. **Relational metadata** is one of the most important parts to be able to explicitly define relations between the content and other related information / content. Relations can include explicit ones which are given through the design of the content object

or external objects. However it might also contain implicit relations gathered by observations or the usage history.

6. **Rights metadata** are of utmost importance with respect to reusability as they declare the terms of use.
7. **Functional metadata**: Functions may be supported to alter the presentation of the content, to customize or personalize the content, or to provide access to different versions.

The selection of the different metadata types with respect to reuse was based on a set of expert interviews and is currently empirically validated in a survey. The assignment of metadata to resources is not restricted to the above mentioned types but is open to domain-specific assignments like educational value for a learning object or preservation data for archival information. In this respect our approach is inline with the vision of Resource Profiles as described in [8].

4.3 Ontology Framework

The data and metadata model as explained in the previous sections are implemented using a set of ontologies in order to publish and describe ICOs on the Web. The import-graph of the ontologies used is depicted in Figure 2²: The main ontology is the RICO

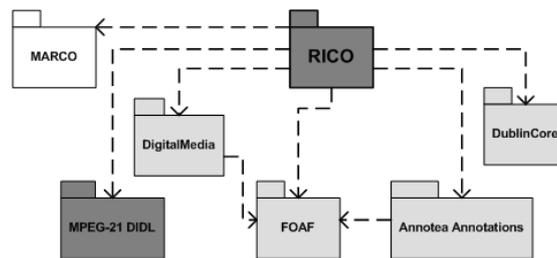


Fig. 2. The Reusable Intelligent Content Objects (RICO) ontology

(“Reusable Intelligent Content Objects”) - ontology which imports a set of other ontologies. RICO is an OWL-DL ontology. Most notably it makes use of

- the MPEG-21 DIDL ontology which we built to reflect the data model presented in section 4.1,
- the Mindswap Digital Media Ontology which is used to type resources³
- an OWL-DL version of the FOAF ontology as provided by Mindswap⁴,

² The dark grey and white ontologies were built in the course of this work.

³ The Digital Media Ontology available at <http://www.mindswap.org/2005/owl/digital-media> has been slightly adapted to be in OWL-DL.

⁴ <http://www.mindswap.org/2003/owl/foaf>

- the Annotea annotations ontology to represent annotations⁵,
- the OWL-Lite version of the Dublin Core ontology, and
- the MARCO (“Metadata for Reusable intelligent Content”) - ontology which is currently work in progress and which will cover aspects of the metadata model as presented in section 4.2.

5 Deployment of Intelligent Content Objects on the Web using RDFa

ICOs are published as compound objects on the (Semantic) Web following the data model described in section 4.1, including metadata as described in section 4.2 and marked up with RDFa using the ontologies as described in section 4.3. An ICO includes the resource whose structure is described using the data model and multiple metadata records including the metadata types previously presented.

The compound package information is about to be published inline within an HTML page and will extend the ramm.x (“RDFa based multimedia metadata”) model that we suggested in [11] and which can be used to deploy multimedia metadata using RDFa.

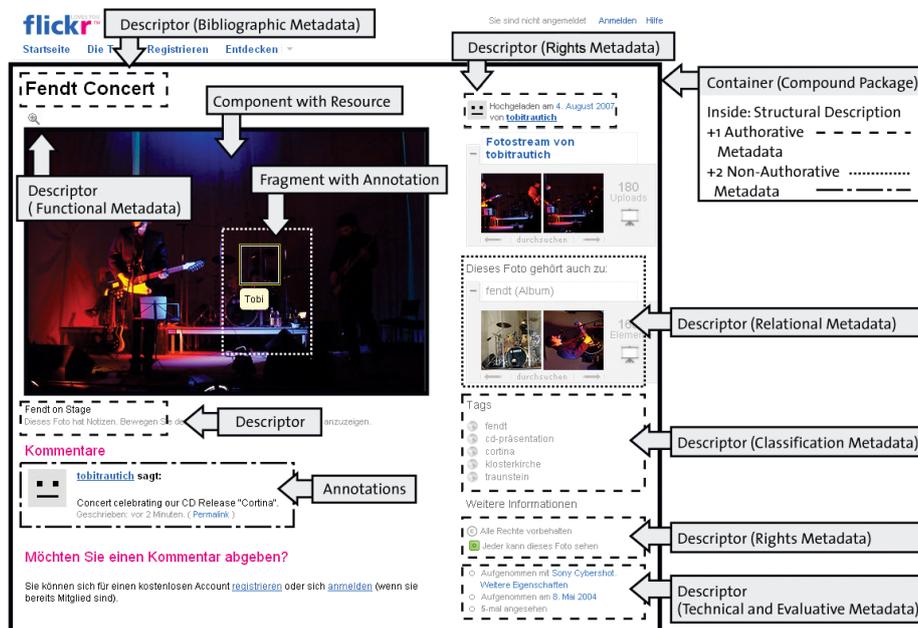


Fig. 3. A image hosted by Flickr deployed as an Intelligent Content package

⁵ The Annotea annotations ontology available at <http://www.w3.org/2000/10/annotation-ns#> has been rebuilt in OWL-DL.

Figure 3 shows an example of an ICO, i.e. an image hosted by Flickr for which all descriptions are explicitly marked up and related to the image. The ICO contains three metadata sets: one authoritative set (as provided by the owner/creator of the image) and two non-authoritative sets (one provided by the hosting platform and one provided by an end user through a commentary). The figure shows only how visible information is related to the image. However further additional (non-visible) metadata could also be provided, e.g. by providing a detailed description of different scenes in a video or further semantic descriptions of the content of an image. Parts of the resulting RDF graph are visualized in Figure 4 which is however not showing the entire graph because of space restrictions (i.e. most descriptors are omitted).

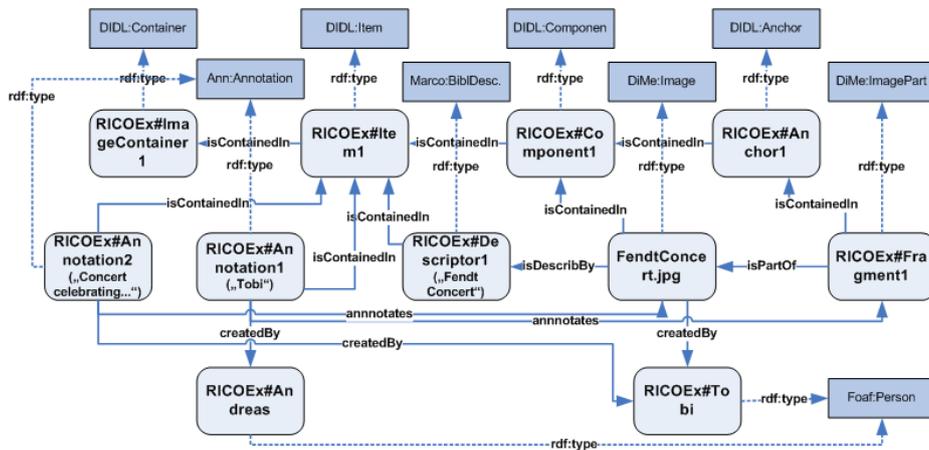


Fig. 4. Partial description of the compound object from Figure 3.

6 Related Work

The model and the ontologies can be used to markup illustrative images in typical web pages, images or videos hosted by social media sharing sites, slides and images embedded in commercial offerings, or multimedia content deployed in blogs and wikis.

Especially in recent years much work has been done on the specification of ontologies that aim to combine traditional multimedia description models, thus trying to develop models that allow reasoning over the structure and semantics of multimedia data (see [5] and [12] for a comprehensive overview). Models with a similar intention, i.e. to publish semantic metadata inline of HTML pages, include lightweight approaches like the hMedia microformat⁶ which is a basic vocabulary to mark up media resources on web sites using property value pairs. Furthermore it is related to the ramm.x model which provides a small but extensible vocabulary for marking up resources to include legacy

⁶ http://wiki.digitalbazaar.com/en/Media_Info_Microformat

metadata. Ramm.x however does not include something similar to our data model or a detailed metadata model. The intention of the SMIL MetaInformation-Module⁷ is to publish RDF-based metadata in SMIL presentations. It is very general and does not prescribe how to use it. We intend to test the applicability of our proposed model with SMIL in the future. The intention of our model is also similar to Adobe's XMP whose intention is to publish RDF-based metadata into PDFs or other document formats. Furthermore we want to acknowledge the work being done by Creative Commons to describe and embed licensing data using RDF which is exploited in searches by Yahoo or in Flickr⁸.

More heavyweight approaches include *Intelligent Content* models as previously assessed for example in [6] and which cover a broad range of aspects. Most of these approaches are too heavy for our proposed model.

Traditional models include the standardized framework of MPEG-7 [15] or packaging formats from the archival or E-Learning domain which include the IMS Content Packaging format⁹, the Metadata Encoding and Transmission Standard (METS)¹⁰, the MPEG-21 Digital Item Declaration (DID) [4] and most recently the OAI-Object Reuse and Exchange - model (OAI-ORE)¹¹. The OAI-ORE model has been designed as an exchange format for scholarly works. Its compound objects model¹² [21] is similar to our model as it also provides facilities to publish semantic descriptions as an overlay graph over web pages. The approach however does not focus on multimedial aspects.

The intention of the presented model is not provide a new standard for the description of the semantic content and content decomposition like it is done by MPEG-7. Thus our approach is only marginally related to endeavors that aim to combine MPEG-7 with semantic technologies like the COMM ontology or other available MPEG-7 ontologies (see cf. [20] for a comparison).

7 Conclusions and Future Work

In this paper we presented a model for deploying multimedia content descriptions, i.e. Intelligent Content Objects, on the Semantic Web with the goal to increase the reusability potential of content in general. The model consists of a data model that supports adaptability of content, a metadata model including properties to explicitly increase findability with respect to reuse (e.g. implicit usage information, ratings, etc.) and a deployment facility to publish content descriptions inline of HTML pages. Deploying resources using this model can have a similar effect like Yahoo's Search Monkey¹³ which allows people to mark up their content using Microformats or RDFa whereas the additional information is then used to display and probably rank search results.

⁷ <http://www.w3.org/TR/2007/WD-SMIL3-20070713/smil-metadata.html>

⁸ <http://search.creativecommons.org/>

⁹ <http://www.imsglobal.org/content/packaging/>

¹⁰ <http://www.loc.gov/standards/mets/>

¹¹ <http://www.openarchives.org/ore/>

¹² <http://www.openarchives.org/ore/documents/CompoundObjects-200705.html>

¹³ <http://developer.yahoo.com/searchmonkey/>

Future work includes the engineering of the MARCO ontology and a qualitative evaluation of our approach. The evaluation of the parts of the model is work in progress. The basic model, which includes the data model and the metadata categorization, meets the requirements of a typical multimedia publishing scenario on the Web and fulfills the criteria of an adaptable data model as defined in [3]. However the effect of the different metadata types on the reuse of content has yet to be validated. We are currently empirically validating the influence of different metadata types on the reuse of content in a study which will be accompanied by an implementation. Different aspect that would also demand special attention but which are beyond the scope of our work include the assessment of the heterogeneity or the consistency of different metadata sets.

References

1. Wernher Behrendt, Guntram Geser, and Andrea Mulrenin. Ep2010 - the future of electronic publishing towards 2010. Information Society DG - Unit E2, EUFO 1-275, Rue Alcide de Gasperi, L-2920, Luxembourg, 2003.
2. Jeroen Bekaert, Emiel De Kooning, and Herbert van de Sompel. Representing digital assets using mpeg-21 digital item declaration. *International Journal on Digital Libraries*, 6(2):159–173, 2006.
3. Susanne Boll and Wolfgang Klas. Zyx – a multimedia document model for reuse and adaptation of multimedia content. *IEEE Transactions on Knowledge and Data Engineering*, 13(3):361–382, 2001.
4. Jan Bormans and Keith Hill. Mpeg-21 overview v.5, 2002. ISO/IEC JTC1/SC29/WG11.
5. Tobias Bürger and Michael Hausenblas. Why real-world multimedia assets fail to enter the semantic web. In *Proc. of the Int. Workshop on Semantic Authoring, Annotation and Knowledge Markup (SAAKM07)*, 2007.
6. Tobias Bürger, Ioan Toma, Omair Shafiq, and Daniel Dögl. State of the art in sws, grid computing and intelligent content objects - can they meet? GRISINO Deliverable D1.1, <http://www.grisino.at>, October 2006. GRISINO Deliverable D1.1.
7. CCSDS. Reference model for an open archival information system. Blue Book 1, Consultative Committee for Space Data Systems, CCSDS Secretariat Program Integration Division (Code M-3) National Aeronautics and Space Administration Washington, DC 20546, USA, Januar 2002.
8. Stephen Downes. Resource profiles. *Journal of Interactive Media in Education Special Issue on the Educational Semantic Web*, 5:0, 2004.
9. Jane Greenberg. Metadata generation: Processes, people, and tools. *Bulletin of American Society for Information Science and Technology*, 29(2), 2003.
10. Lynda Hardman. Canonical processes of media production. In *MHC '05: Proceedings of the ACM workshop on Multimedia for human communication*, pages 1–6, New York, NY, USA, 2005. ACM Press.
11. Michael Hausenblas, Werner Bailer, Tobias Bürger, and Raphael Troncy. Ramm.x: Deploying multimedia metadata on the semantic web. In *Proceedings of SAMT 2007, Dec 4-7, Genova, Italy, 2007*.
12. Michael Hausenblas, Susanne Boll, Tobias Bürger, Oscar Celma, Christian Halaschek-Wiener, Erik Mannens, and Raphael Troncy. Multimedia vocabularies on the semantic web. W3c incubator group report, W3C, July 2007.
13. Laura Hollink, Guus Schreiber, Bob Wielinga, and Marcel Worring. Classification of user image descriptions. *Int. J. Hum.-Comput. Stud.*, 61(5):601–626, 2004.

14. Harald Kosch, Laszlo Boszormenyi, Mario Doller, Mulugeta Libsie, Peter Schojer, and Andrea Kofler. The life cycle of multimedia metadata. *IEEE MultiMedia*, 12(1):80–86, 2005.
15. José M. Martínez-Sánchez, Rob Koenen, and Fernando Pereira. Mpeg-7: The generic multimedia content description standard, part 1. *IEEE MultiMedia*, 9(2):78–87, 2002.
16. Frank Nack, Jacco van Ossenbruggen, and Lynda Hardman. That Obscure Object of Desire: Multimedia Metadata on the Web (Part II). *IEEE Multimedia*, 12(1), 2005.
17. Mimi Recker and David Wiley. A non-authoritative educational metadata ontology for filtering and recommending learning objects. *Journal of Interactive Learning Environments*, 2001.
18. Ann Rockley. *Managing Enterprise Content: A Unified Content Strategy*. New Riders, 2002.
19. Arnold W. M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, and Ramesh Jain. Content-based image retrieval at the end of the early years. *IEEE Trans. Pattern Anal. Mach. Intell.*, 22(12):1349–1380, 2000.
20. Raphael Troncy, Oscar Celma, Suzanne Little, Roberto Garcia, and Chrisa Tsinaraki. Mpeg-7 based multimedia ontologies: Interoperability support or interoperability issue? In *Proceedings of the 1st Workshop on Multimedia Annotation and Retrieval enabled by Shared Ontologies (Mareso '07)*, 2007.
21. Herbert van de Sompel and Carl Lagoze. Interoperability for the discovery, use, and re-use of units of scholarly communication. *CTWatch Quarterly*, 3(3), August 2007.
22. Jacco van Ossenbruggen, Frank Nack, and Lynda Hardman. That Obscure Object of Desire: Multimedia Metadata on the Web (Part I). *IEEE Multimedia*, 11(4), 2004.