Multimedia Metadata Standards

Werner Bailer, Lionel Brunie, Mario Döller, Michael Granitzer, Ralf Klamma, Harald Kosch, Mathias Lux, Marc Spaniol

Multimedia Metadata Standards

Werner Bailer[1], Lionel Brunie[2], Mario Döller[3], Michael Granitzer[4], Ralf Klamma[5], Harald Kosch[3], Mathias Lux[6], Marc Spaniol[7]
[1]Joanneum Research, Graz, Austria
[2]LIRIS, INSA de Lyon, Villeurbanne, France
[3]University of Passau, Passau, Germany
[4]Know-Center Graz, Austria
[5]RWTH Aachen, Aachen, Germany
[6]Klagenfurt University, Klagenfurt, Austria
[7]Lehrstuhl Informatik, Aachen, Germany

Synonyms

► Metadata standards; ► Metadata model; ► Metadata semantics; ► Metadata validation; ► Metadata applications

Definition

Multimedia metadata describes various aspects of multimedia content, including formal and technical properties (e.g. encoding, format), information about the creation of content, the processing applied, its use, rights information and the structure and semantics of the content itself. The importance of these aspects depends on the role of the provider or consumer of the metadata and the stage in the media life cycle. Multimedia metadata is among others important for production, distribution, search and indexing of multimedia content.

Introduction

Multimedia metadata is data describing various aspects of multimedia data. In general, one can distinguish between *low-level* and *high-level* multimedia metadata. Low-level metadata can be extracted automatically and describes characteristics like color, texture or shape in case of visual information or like timbre, pitch and rhythm in case of audio information. High-level metadata on the other hand allows descriptions on a conceptual level. This includes the specification of digital rights, content summaries or related content. In opposite to low-level metadata, the creation of high-level metadata requires manual input by human annotators. The inability to create high-level metadata in a fully automated way is called *semantic gap* or *sensory gap* (see [1]).

Multimedia metadata standards in general formalize a *metadata model*, which describes the semantics and value restrictions of description elements as well as relations between description elements. Using this model metadata can be validated. Elements (or sets of elements) of a metadata model describe a certain aspect of multimedia content. Such a description is also called *feature*. Depending on whether a feature belongs to low- or high-level metadata one can speak of low-level and high-level features, respectively. Features of the same type can be compared to each other by computing a pair wise distance or similarity. Comparisons like this are employed in metadata based content search and retrieval, indexing and visualization. A common example for a low-level feature is a color histogram, where the color characteristics are described by a vector. An appropriate distance function would be for instance the angle between two vectors or the L_1 distance. A high-level feature would be a genre classification, whereas the distance function can be defined for instance on a taxonomy.

In the next section we describe important multimedia metadata standards currently in use (P_Meta, SMPTE-DMS-1, DC, MPEG-7, and TVA). The following section introduces to metadata semantics. Then metadata validation is discussed. Finally important metadata applications are pointed out and a conclusion is given.

Metadata Standards

P_Meta – EBU P_Meta Project The P_Meta Semantic Metadata Schema is the result of P_Meta Project of the European Broadcasting Union (EBU). It aims to provide a common vocabulary for exchange of information capturing program in professional broadcasting industry. P_Meta provides disambiguated definitions of metadata elements in human language concerning different aspects of descriptive metadata. P_Meta is based on XML and is a royalty free standard. P_Meta applies primarily to Business-to-Business (B2B) scenarios including roles of content creators, who create new material, and content distributors, who publish and deliver material, as well as content archives, which preserve material and enable re-use.

The P_Meta Metadata Model defines Programmes, which are organized in Programme Groups. The Programmes include MediaObjects, Items and ItemGroups. ItemGroups are collections of Items, which are fragments of a Programme. A MediaObject is continuous in time and a single component of a Programme or Item. All actual description elements as well as the structure are defined by XML elements, whereas possible values are given in classification schemes. For example an element describing a relationship between different Materials is named "MaterialRelationshipTypeCode" and has possible values defined in the "tva:HowRelatedCS" classification scheme from TV Anytime. Possible values include "Trailer," "IsMakingOfFor" or "Promotional Still Image," which are expressed in numbers as defined in the classification scheme. A full list of elements including links to the respective classification schemes can be found in [2].

SMPTE Metadata Dictionary and DMS-1 The SMPTE Metadata Dictionary is standardized by the Society of Motion Picture and Television Engineers (SMPTE). Instead of defining an extensive model for metadata, it is mainly a large list of structured metadata elements. For these elements keys, value length and semantics are defined in the metadata dictionary. The current specification (see [3]) defines a tree structure with around 500 nodes and 1,500 leafs describing elements in 6 distinct classes: identification, administration, interpretation, parametric, process, relational and spatio-temporal. For each element, a 16 byte key is given, comprising the unique identifier of the element as well as the version number of the dictionary reflecting when it was introduced.

Like P_META, the Metadata Dictionary applies generally to B2B scenarios and is used in conjunction with the Material Exchange Format (MXF). MXF combines the essence (actual multimedia content like audio or video streams) with metadata for exchange between different parties. Metadata is integrated through Descriptive Metadata Schemes (DMS) and DMS-1 defines the use of the SMPTE Metadata Dictionary in MXF.

Dublin Core Metadata Initiative The Dublin Core Metadata Initiative (DCMI) [4,5], an open organization engaged in the development of interoperable online metadata standards founded in 1995, has developed the Dublin Core (DC) Metadata Standard. DC targets at the high-level description of networked resources like video, audio, images and text as well as their multimodal combination into documents. It is a rather simple standard compared to for example P_Meta, SMPTE Metadata Dictionary and MPEG-7 since networked and internet resources are in its primary focus, DC may be viewed as rather simple predecessor of RDF (the Resource Description Framework).

For describing a resource Dublin Core elements are assigned to the particular resource, whereas one element may be assigned several times or never. DC Elements can be seen as simple key-value pairs, where keys are drawn from a small, predefined set of element types and values are filled in accordingly.

Two different levels of description details are considered in the DC standard version 1.1: simple and qualified. Simple Dublin Core consists of 15 elements like for example Title, Creator, Publisher, Subject etc. Those 15 elements cover viewpoints of time, persons, content, format as well as rights on a single resource and have been endorsed in the NISO Standard Z39.85–2001 and ISO Standard 15836–2003.

The Qualified Dublin Core level introduces three additional elements ("audience," "provenance" and "rightsholder") as well as refinements (also called qualifiers). Those refinements allow for semantically refining the meaning of the Simple Dublin Core elements, whereas they do not provide extension mechanisms for new elements. One example is the element-refinement "abstract," which refines the basic element "description," While the "description" element may contain table of contents, images or free text annotation, the refined "abstract" element must contain "a summary of the content resource." Additionally to the semantic enrichment of elements, refinements provide means to describe the coding scheme for one element. For example, for the element "identifier" one can specify, whether the element is encoded as URI (Uniform Resource Identifier) or in another format.

MPEG-7 MPEG-7, formally known as ISO/IEC 15938 – Multimedia Content Description Interface, is the ISO/IEC international multimedia description standard developed by MPEG (Moving Picture Experts Group). MPEG-7 (latest version in 2006) concentrates on describing multimedia content in a semantically rich manner. While earlier MPEG versions were focusing on making multimedia content available, MPEG-7 allows the localization, filtering, managing and processing of desired media files by tagging them with information regarding their content and origin (see [6,7]).

From a technical point of view, MPEG-7 standardizes *Descriptors (D)* for defining the syntax and semantics of each feature representation and *Description Schemes (DS)* for specifying relationships between components (both descriptors and description schemes). Both, DS and D, can be defined and modified with the help of the *Description Definition Language (DDL)* which is based on XML Schema extended by new data types, like for feature vector representation. In addition, Classification Schemes (CS) support the extensibility of MPEG-7. This includes parental rating, and content classification into a number of pre-defined categories. MPEG-7 is organized in the following:

MPEG-7 Systems: It includes tools for preparing MPEG-7 Descriptions to allow an efficient transport and storage (BiM – binary format for MPEG-7). Further, there are mechanism for supporting synchronization between content and their descriptions. Additionally MPEG-7 Systems provide tools for managing and protecting intellectual property.

MPEG-7 Description Definition Language: The DDL is one of the core parts that are used for instantiating MPEG-7 descriptions. It provides a descriptive foundation for creating user defined description schemes and descriptions.

MPEG-7 Visual: These description tools consist of basic description schemes and descriptors that mainly cover color, texture, shape, motion and localization. For describing an image, one has for instance the possibility to define the used color space or the dominant color.

MPEG-7 Audio: These description tools comprise basic description schemes and descriptions that cover basic audio features: sound effect description tools, melodic descriptors for query-by-humming and spoken content description which can be classified as high-level descriptors and a audio description framework for describing low-level features (e.g., AudioSignatureTpye).

MPEG-7 Multimedia Description Schemes (MDS): They standardize on the one hand generic descriptors and descriptor schemes and on the other hand multimedia entities that represent more complex structures. Generic features are basically descriptors that are common to all media (audio, visual and text) for instance, vector, time, etc. Apart from standardizing generic features, MDS provides more complex description tools such as Content Management, Content Description, Navigation and Access, Content Organization and User Interaction.

MPEG-7 Reference Software: It is a platform for testing all corresponding Descriptors (D), Descriptor Schemes (DS), etc.

MPEG-7 Conformance: It defines guidelines and approaches for testing descriptions and processing engines.

MPEG-7 Extraction and Use of Descriptions: It provides information about using some of the description tools. It can be seen as an additional part to the Reference Software.

MPEG-7 Profiling: It provides the possibility to restrict descriptions of the MPEG-7 schema. Detailed information is presented in the adjacent article "Multimedia Metadata Profiles."

MPEG-7 Schema Definition: This part contains means (e.g., description tools, namespace designator) for creating different versions of the MPEG-7 schema using the Description Definition Language (DDL).

MPEG-7 Profile Schemas: This part demonstrates several MPEG-7 based profiles.

MPEG Query Format: It is the latest one and is currently in FCD status. It standardizes the access to multimedia retrieval systems.

TV-Anytime – TV-Anytime Metadata TV-Anytime (TVA) Metadata has been designed to support the Business-to-Consumer exchange in the broadcast industry [8,9].

It allows the consumer to find, navigate and manage content from a variety of internal and external sources including, for example, enhanced broadcast, interactive TV, Internet and local storage. Metadata is generated during the process of content creation and content delivery. There are three basic kinds of metadata: *Content Description*, *Instance Description*, and *Consumer Metadata*. In addition, the standard defines *Segmentation Metadata* and *Metadata Origination Information Metadata*. The information that the consumer or agent will use to decide whether or not to acquire a particular piece of content is called *attractors*, and is used in electronic program guides, or in Web pages. These attractors rely on descriptors stemming from MPEG-7. Furthermore, some MPEG-7 datatypes are used directly (e.g., mpeg7:TextualType is used for many TVA of elements).

The *content description metadata* describes content independently of any particular instantiation of a media programme. Programme in this context means an editorially coherent piece of content. Descriptions of content, e.g., television programs are held in the *ProgramInformationTable*. They include metadata like the Title (here the mpeg7:TitleType is used) of the program, a Synopsis, the Genre it belongs to, and a list of Keywords that can be used to match a search. Descriptions of groups of related items of content e.g. all episodes of "Foxes in the Wild" are held in the *GroupInformationTable*. They include among other the GroupType, a BasicDescription and MemberOf element. A mapping of cast members to unique identifiers is held in the *CreditsInformationTable*. The identifiers can be used in other metadata instances simplifying the search. The purchase information, like Price and PurchaseType, is held in the *PurchaseInformationTable*. Critical reviews of items of content are held in the *ProgramReviewTable*. They include metadata like the Reviewer, a FreeTextReview and the ProgramId.

Instance Description Metadata is required in case of significant differences between instantiations of the same content. These are instances with the same Content Reference Identifier (CRID) (the CRID connects content metadata with content). Instance Metadata is connected with content related to a definite event. Descriptions of particular instances (locations) of content are held in the *ProgramLocationTable*. They include the elements Schedule, BroadcastEvent, OnDemandProgram and OnDemandService, all derived from ProgramLocationType. They include among other information about the program, start and end. Also Title, Synopsis, Genre and PurchaseList can be specified. Descriptions of services within a system are held in the *ServiceInformationTable*. For each single service, the Name, Owner Logo (here the mpeg7:MediaLocatorType is used), ServiceDescription, ServiceGenre etc. can be specified.

Consumer Metadata includes Usage History and User Preferences, both based on respective MPEG-7 datatypes. The Usage History provides a list of the actions carried out by the user over an observation period. It is used for tracking and monitoring the content viewed by individual members. Thus, it builds a personalized TV guide by tracking user viewing habits, selling viewing history to advertisers or tracking and monitoring content usage for more efficient content development. The User Preferences facilitate description of user's preferences pertaining to consumption of multimedia material. They include FilteringAndSearchPreferences and BrowsingPreferences and can be correlated with media descriptions to search, filter, select and consume desired content.

Metadata Semantics

When speaking about semantics in general and metadata (data about data) semantics in particular it is worth to mention that the meaning of any data might only be disclosed based on its situational context. However, metadata have to be created, recorded, stored, processed, transferred, distributed and deployed without prior knowledge of the situational context, like the application, user community, time, space etc. For this reason, metadata have to be embedded into the context of metadata schema specification like MPEG-7 or into metadata models which add the necessary context information for metadata semantics by the work of international standardization committees or by the effort of modeling.

In the domain of multimedia processing we observe a semantic gap between the technical extraction of metadata and the semantically correct interpretation of content. In the case of multimedia we distinguish between content-dependent and content-descriptive metadata [10]. Content-dependent metadata consist of so-called low-level features and are automatically recorded and stored with the multimedia content. A typical example is the Exif [11] metadata which are recorded and stored in modern digital cameras, even with GPS [12] metadata in case a receiver is available. Furthermore, Content-dependent metadata can be extracted from the multimedia content itself. A typical example for automatic extraction and processing of low-level features is the color layout feature vector of an image expressed by a numerical vector (64 coefficients). As an application two images (here a sketch and an image) can now be compared by computing the distance between the two vectors. The metadata here can be obtained and processed without any user interaction. Content-descriptive metadata are typically created manually through users using an application, like annotating (tagging) images with keywords in Flickr or providing free-text descriptions. These annotations and descriptions are considered to carry high-level semantics.

The semantic problem which is often described in the literature is on the one hand, that metadata automatically recorded, extracted and processed are not really helpful in multimedia metadata applications like multimedia retrieval because the results do not match the expectation of users. On the other hand, annotations and descriptions were a long time considered as problematic because users were assumed not to volunteer delivering them. Obviously, the latter problem is eased by a lot of multimedia application becoming web-based where huge user numbers deliver metadata for multimedia content. However, the quality of the metadata can not be guaranteed by any means. On the basis of the previous illustration it is quite natural that the semantics of any multimedia content can be improved by considering both the content-dependent metadata and the content-descriptive metadata.

In the following, we will introduce both concepts, with a focus on standard compliant information processing. Being an easy to understand and concise method for media annotations, the Dublin Core metadata standard has been a step forward into that direction. In this aspect, Dublin Core has become quite popular in a somewhat coarse (tagging alike) annotation of high-level semantics in the domain of print media. However, Dublin Core is not well suited for temporal and media specific annotations of multimedia contents. For that reason, multimedia metadata standards like MPEG-7 and MPEG-21 overcome these limitations by more sophisticated description elements for time based media. Even more, it is possible to define mappings between MPEG-7 and less sophisticated metadata standards like Dublin Core [13,14] or Cidoc CRM [15]. We will stick to MPEG-7 because it offers the semantically richest metadata model for the description of multimedia semantics.

The so-called Web 2.0 has yet not fully turned into a multimedia web but impressive numbers of metadata have been created for multimedia content, enough to become a research topic on its own. With the opportunities given by the combination of metadata descriptions standards like MPEG-7 and Web 2.0 like participation, the manageability of multimedia semantics in user communities becomes more feasible: The semantics of a multimedia content is the result from the set of annotations or descriptions created or linked in the user community. In contrast to available approaches the reference on MPEG-7 provides a standardized metadata vocabulary, thus ensuring exchangeability and understandability of multimedia contents across applications and domains. Even more, MPEG-7 has advanced features to describe and manage multimedia artifacts as well as multimedia collections. Hence, using MPEG-7 to capture content-dependent and content-descriptive metadata allows the community members to use these information for browsing multimedia artifacts, searching multimedia collections, and navigating hypermedia graphs.

Metadata Validation

Validation of metadata descriptions means to check the conformance and consistency of a metadata document with respect to a metadata standard (or a profile of a standard). This is for example necessary to ensure the correctness of descriptions produced by a system and for publishing and exchanging metadata. Validating metadata documents is also a prerequisite for converting or mapping them to different formats or standards.

As the definition of a metadata standard typically encompasses several layers of representation, validation has to consider each of these layers. The lowest layer is in most cases the encoding of the metadata document[1], for example, XML or KLV[2]. The next layer is a syntactic definition of the elements of the description, such as an XML DTD, XML Schema or the set of

[1] Some standards do not specify this layer.

[2] KLV (Key-Length-Value) encodes items into Key-Length-Value triplets, where key identifies the data or metadata item, length specifies the length of the data block, and value is the data itself [16].

keys of a KLV-based representation. The top layer includes the semantics of each of the description elements, their context and their relation.

There exist tools for validating a metadata document w.r.t. the two lower layers. For example, in the case of a XML Schema based representation (such as supported by MPEG-7, TV-Anytime, P_Meta, etc.), standard XML parsers and XML Schema validators can be used to check first the well-formedness of the XML document and then the validity w.r.t. the schema.

Validating a metadata document in terms of semantics is much more difficult. The first issue is the definition of the semantics of the description elements. In some standards, this definition is only very general and fuzzy, for example, it is not specified whether the Dublin Core element "Title" is meant to contain the working title of the program, the title in which it has been published in a certain language, the title of a series or that of one episode, etc. In those standards where there is a clear definition of the semantics of the elements (e.g. P_Meta, MPEG-7, TV-Anytime) it is in textual form. In addition, correct "golden files" are provided as reference. This makes it possible for a human to check the validity of a document, but the lack of a formalization of the semantics makes automatic validation impossible.

The validation of description semantics can be built into an application. However, this "hard-wires" the semantics into the application's code and does not allow for easy exchange and modification of the semantics of a description. An alternative approach for the semantic validation of documents w.r.t. MPEG-7 profiles has been proposed in [17]. The work is based on MPEG-7 profiles (see short article on metadata profiles), as the semantics of description tools can be better defined in the context of a profile than for the general standard. The semantic constraints defined by a profile are modeled using an ontology and rules. The ontology only models constraints which cannot be expressed in XML Schema. The MPEG-7 document to be checked is transformed to RDF using a XSL stylesheet based on the ontology, and reasoning and rule-checking are performed to detect violations of the semantics.

Once a formalization of the semantics of a standard's description elements is available, it can serve as a basis for defining mappings between different standards based on mappings between the ontologies. The advantage of this approach is that syntactic and modeling differences can be partly abstracted.

Metadata Applications

While in multimedia applications for image or video retrieval previously multimedia processing capabilities were elaborated but contents were the problem, in the Web 2.0 the contents are not a problem anymore, but the metadata quality and the limited processing capabilities of web-based, browser-based or even mobile applications are problematic.

Social Software is a good example for such new applications. Social software concentrates on the link-up between social entities in digital social networks and their interaction, typically by sharing multimedia content. Software is mainly realized by means of computer-mediated communication. It simply turns users into content prosumers (consumer and producer in parallel), anytime and anywhere. Tons of multimedia artifacts are created in the many social software applications available like flickr for image sharing youtube for video sharing, last.fm for audio sharing. As a result, the spirit of the Web 2.0 has resulted in a great amount of multimedia artifacts and a large number of diverse user communities linked a huge so-called social graph.

The context of a multimedia content can be considered neither static nor universally valid. The main problem is metadata quality or uncertainty. On the low-level, it addresses automatic capturing, storing, processing, distributing, etc. of multimedia content on thousands of different devices like mobile phones, digital cameras, radars and satellites. On the high-level, it concerns the different interpretation and understanding of content by users based on their cultural, intellectual and societal background.

While we can exchange multimedia data among different social software applications, metadata are sticky, because they are so invaluable presenting the real currency of the Web 2.0. Therefore, social software is mostly incapable of reusing metadata created by another application, as they are not capable of understanding the other application's metadata. While recently, mash-ups appear to become the missing link, a closer look proves that they are not solving the problem on a large scale but only on a bi-application level. With the draw on metadata standards it becomes possible to gain maturity in Social Software and to reuse metadata not only in diverse applications, but also across different communities. Hence, we will introduce metadata applications based on the MPEG-7 metadata standard functioning on different semantic levels.

Low-Level Semantic Multimedia Processing Low-level semantic multimedia processing mostly deals with features automatically generated from the multimedia content. These are low-level numerical feature vectors of audio-visual contents. For this purpose, MPEG-7 offers an extensive metadata model covering a whole range of aspects (e.g., production, distribution, storage, rights, transmission, usage). It provides a rich set of description schemata, which allow describing the content in structural (space and time) as well as semantic aspects. A sample application for complex querying based on semantic descriptors in MPEG-7 is Caliph & Emir, a photo annotation and retrieval tool [18,19].

High-Level Semantic Multimedia Processing In order to overcome the problems with interpreting semantics of audio-visual contents correctly high-level metadata in the form of annotations are used. The MPEG-7 standard provides dedicated descriptors for high-level metadata annotations in order to describe the content's semantic aspects. These annotations reach from textual content descriptions up to domain specific "ontologies" of multimedia collections. A prototypical application is MECCA, a multi-dimensional multimedia screening and classification platform [20]. Thus, it serves as a media classification and monitoring system in multimedia-centric, interdisciplinary knowledge exchange processes. The MPEG-7 Encoding of Dublin Core Information and Naming Application (MEDINA) has been developed in order to support collaboration in communities by the exchange of multimedia contents and high-level semantic descriptions [14]. Thus, MEDINA is based on an excerpt of the extensive MPEG-7 multimedia metadata standard with an integrated semi-automatic Dublin Core to MPEG-7 conversion functionality. Taking into account Web 2.0 paradigms such as collaborative tagging and the long tail of specific user communities, the Nillenposse Media Viewer is a cross-media and cross-community high-level annotation tool [21]. Therefore, it supports community-aware semantic multimedia tagging by so-called Commsonomies (a combination of Communities and Folksonomies).

Hybrid Semantic Multimedia Processing In order to bridge the gap between high-level semantic information about multimedia contents and their purely technical low-level content descriptions, a hybrid approach – combining content-dependent and content-descriptive metadata – is required. Because of this being by far more complex than the stand alone solutions introduced before, these applications are quite rare. A sample application is Imagesemantics [22]. It is a concise image retrieval system that allows the retrieval of images by combining low-level content based features and high-level metadata annotations. Another approach aiming at the exploitation of the Web 2.0 in a comprehensive community information system is Virtual Campfire [23]. It is a Social Software that allows a modular composition of web services based on its lightweight application server. Thus, it allows hybrid semantic multimedia processing covering geospatial content sharing, multimedia tagging and collaborative authoring of multimedia contents [24].

Conclusion

The field of multimedia metadata is diverse in multiple aspects. Intentions and constraints differ in multiple aspects based on different scenarios (B2B, B2C) and different roles (producer, distributor and consumer). Complexity depends on the trade off between simplicity and comprehensiveness as well as value restrictions (e.g. by taxonomies) and interoperability. Standardization groups and organizations like EBU P/MAG, SMPTE and MPEG overlap in topics and interests and therefore several standardized formats describing same aspect in different ways exist.

Consumers on the other side do not participate in the efforts put in standardization and interoperability: While consumption of metadata by using the Electronic Programme Guide (EPG) or DVD sub titles is common; annotation by the home user is currently not discussed at large. Therefore several de facto standards in this area, like ID3 for the annotation of audio files in personal digital libraries, have emerged. De jure standards like MPEG-7 on the other hand have virtually no impact on personal digital libraries.

It can be assumed that there is a long way to go in the area of multimedia metadata standards in multiple aspects: The most prominent topic is how to *bridge the semantic gap*. But nonetheless important is the question how to *harmonize different standards* and allow interoperability between different applications and domains. Last but not least an open question is how to allow every element in the multimedia production and consumption lifecycle to re-use and extend metadata and ensure *timeliness and relevance* concerning the multimedia essence.

References

1. S. Santini and R. Jain, "Beyond Query By Example" Proceedings of the Sixth ACM International Conference on Multimedia, pp. 345–350, Bristol, United Kingdom, 1998.
2. P_Meta 2.0 Metadata Library. EBU-TECH 3295-v2, 2007.
3. Metadata Dictionary Registry of Metadata Element Descriptions. SMPTE RP210.10, 2007.
4. Dublin Core Metadata Initiative, http://dublincore.org/, visited 12/2007.
5. Dublin Core Usage Guidelines, http://dublincore.org/documents/usageguide/visited 12/2007.
6. H. Kosch, L. Böszörményi, M. Döller, A. Kofler, P. Schojer, and M.Libsie, "The life cycle of multimedia meta-data," IEEE MultiMedia, Vol. 12, No. 1, January/March 2005.
7. B.L. Tseng, C.-Y. Lin, J.R. Smith, "Using MPEG-7 and MPEG-21 for Personalized Video," IEEE Multimedia, 11(1), pp. 42–52, 2004.
8. ETSI TS 102 822-3-1 V1.3.1 Technical Specification. Broadcast and online services: Search, select, and rightful use of content on personal storage systems ("tv anytime"); part 3: Metadata; subpart 1: Phase 1 metadata schemas, June 2005.
9. ETSI TS 102 822-2 V1.3.1. Broadcast and online services: Search, select, and rightful use of content on personal storage systems ("tv anytime"); part 2: System description, June 2005.
10. A. Del Bimbo, "Visual Information Retrieval," Morgan Kaufmann, Los Altos, CA, 1999.
11. "Exchangeable Image File Format" (EXIF), http://www.exif.org
12. GPS: "Global Positioning System: Serving the World," http://www.gps.gov/, 2007.
13. J. Hunter: "A proposal for the integration of Dublin Core and MPEG-7," ISO/IEC JTC1/SC29/WG11 M6500, 54th MPEG Meeting, La Baule, 2000.
14. M. Spaniol and R. Klamma, "MEDINA: A Semi-Automatic Dublin Core to MPEG-7 Converter for Collaboration and Knowledge Management. In: Multimedia Repositories," Proceedings of I-KNOW'05, Graz Austria, J.UCS Proceedings, LNCS 1590, Springer, Berlin, pp. 136–144, 2005.
15. J. Hunter, "Combining the CIDOC CRM and the MPEG-7 to describe multimedia in museums," Proceedings of the International Conference about Museums and the Web, Boston, Massachusetts, 2000.
16. Data Encoding Protocol using Key-Length-Value. SMPTE 336M, 2001.
17. R. Troncy, W. Bailer, M. Hausenblas, P. Hofmair, and R. Schlatte, "Enabling Multimedia Metadata Interoperability by Defining Formal Semantics of MPEG-7 Profiles," Proceedings of First International Conference on Semantic and Digital Media Technologies, Athens, GR, December 2006.
18. M. Lux and M. Granitzer, "Retrieval of MPEG-7 based Semantic Descriptions," Proceedings Workshop "WebDB meets IR" being Part of 11th GI-Fachtagung für Datenbanksysteme in Business, Technologie und Web, 2005.
19. M. Lux, W. Klieber, and M. Granitzer, "Caliph and Emir: Semantics in Multimedia Retrieval and Annotation," 19th International CODATA Conference, Berlin, Germany, November 2004.
20. R. Klamma, M. Spaniol, and M. Jarke, "MECCA: Hypermedia Capturing of Collaborative Scientific Discourses about Movies," Informing Science: International Journal of an Emerging Discipline, Vol. 8, 2005, N. Sharda (Ed.), Special Series on Issues in Informing Clients using Multimedia Communications, pp. 3–38.
21. R. Klamma, M. Spaniol, and D. Renzel, "Community-Aware Semantic Multimedia Tagging – From Folksonomies to Commsonomies," K. Tochtermann, H. Maurer, F. Kappe, and A. Scharl (Eds.), Proceedings of I-Media'07, International Conference on New Media Technology and Semantic Systems, Graz, Austria, September 5–7, JUCS (Journal of Universal Computer Science) Proceedings, pp. 163–171, 2007.
22. M. Spaniol, R. Klamma, and M. Lux, "Imagesemantics: User-Generated Metadata, Content Based Retrieval and Beyond," JUKM (Journal of Universal Knowledge Management), Special Issue "Best of I-KNOW '07," Springer, Berlin, 2008.
23. Y. Cao, M. Spaniol, R. Klamma, and D. Renzel, "Virtual Campfire – A Mobile Social Software for Cross-Media Communities," K. Tochtermann, H. Maurer, F. Kappe, and A. Scharl (Eds.), Proceedings of I-Media'07, International Conference on New Media Technology and Semantic Systems, Graz, Austria, September 5–7, JUCS (Journal of Universal Computer Science) Proceedings, pp. 192–195, 2007.
24. R. Klamma, Y. Cao, and M. Spaniol, "Smart Social Software for Mobile Cross-Media Communities," M. Lux, M. Granitzer, and M. Spaniol (Eds.), Multimedia Semantics – The Role of Metadata, Springer, Berlin, ISBN 978–3–540–77472–3, 2008.