Unified Access to Media Metadata on the Web

Florian Stegmaier, Werner Bailer, Tobias Burger, Mari Carmen Suarez-Figueroa, Erik Mannens, Jean-Pierre Evain, Martin Hoffernig, Pierre-Antoine Champin, Harald Kosch, and Mario Doller

In: IEEE Multimedia, 20 (2), 22-29, 2013.

Unified Access to Media Metadata on the Web

Florian Stegmaier
University of Passau, Germany

Werner Bailer
Joanneum Research, Austria

Tobias Bürger
Payback, Germany

Mari Carmen Suárez-Figueroa
Universidad Politécnica de Madrid, Spain

Erik Mannens
iMinds - Ghent University, Belgium

Jean-Pierre Evain
EBU

Martin Höffernig
Joanneum Research, Austria

Pierre-Antoine Champin
Université Claude Bernard Lyon 1, France

Mario Döller and Harald Kosch
University of Passau, Germany

The past few decades have seen great advancements in the area of Web-based information retrieval. The aim of this research is to help users access multimedia metadata documents on the Web in an unified way. Although these systems typically analyze the textual content of Web resources, the availability of sophisticated metadata and the establishment of a semantic understanding of the content are essential for retrieving multimedia resources.[1] To promote effective multimedia retrieval, researchers have attempted to create a multimedia life cycle that covers all the steps between content creation and consumption.[2] The key enabler here has been metadata standards, which let us interchange information about media resources between different stages and life-cycle peers in an interoperable way.[3] This is also highlighted in the work of Lynda Hardman and her colleagues on canonical processes of semantically annotated media production in which metadata is omnipresent and essential for guaranteeing the efficiency and quality of the overall process.[4]

Transitioning content features to metadata to make multimedia content as searchable as text is thus essential for every multimedia information retrieval system. Because the multimedia domain is so broad, multimedia resources are typically indexed using different, divergent standards, which makes retrieving multimedia resources even more difficult.[5] The abundance and diversity of content represented in different formats on the Web thus directly translates into a need for a common framework for describing and representing content by covering essential syntactic and semantic features.

Many efforts have attempted to define a common lingua franca for the retrieval of Web-based multimedia resources, starting with the well-known MPEG-7 standard. However, MPEG-7's adoption is still limited because of its complexity and lack of interoperability. Other metadata standards efforts have been lifted to Semantic Web compatible representations with the goal of improving their interoperability.[6] Subsequent efforts[7,8] led to an increased interoperability and reduced complexity but still have not established a core vocabulary to align metadata formats, a key enabler to establishing syntactic and semantic interoperability. (See related work for a comprehensive survey of interoperability issues.[9])

To reach this goal, the World Wide Web Consortium (W3C) launched the Multimedia Semantics Incubator Group (www.w3.org/2005/Incubator/mmsem) to show the feasibility of using Semantic Web technologies to align different multimedia metadata formats and to provide a scalable solution for multimedia metadata interoperability on the Web. The outcome of this group led to the foundation of the W3C Video on the Web activity (www.w3.org/2008/WebVideo).[10] Among others, this activity hosts the Media Annotation Working

Group (MAWG, www.w3.org/2008/WebVideo/Annotations), which aims to develop a core vocabulary for retrieving multimedia resources on the Web. This core vocabulary is not just another metadata format; it is envisioned as being able to ensure that the information exchange among different metadata formats used by (canonical) processes or systems follow standardized and clear semantics. This standardization effort can be directly injected into the entire multimedia life-cycle process to improve interoperability with respect to data exchange.

The main output of the MAWG is the *Ontology for Media Resources 1.0*.[11] The purpose of the ontology is to overcome the current proliferation of multimedia metadata formats. It addresses the interoperability problem between different metadata formats by providing a common set of properties, agnostic to any representation. It specifies the basic metadata needed for media resources as well as syntactic and semantic links between their values in the existing vocabularies. It is accompanied by the *API for Media Resources 1.0*,[12] which provides uniform access to the elements it defines. Furthermore, the API enables the definition of queries irrespective of the underlying format in which the actual media resource is described.

Requirements

The requirements for the Ontology for Media Resources and API for Media Resources are derived from two sources. The first is the conclusions from the Multimedia Semantics XG that led to the definition of the MAWG charter. The second is the scientific literature discussing multimedia metadata interoperability issues and their potential solution. Finally, the working group collected a set of use cases, derived a set of requirements from them, and distilled a set of core requirements. (See www.w3.org/TR/media-annot-reqs for use cases and requirements.) Thus, both specifications are coined by the following requirements:

Coverage: The ontology and API are not bound to a specific application domain, media type, or content representation.

Composition: The design of the ontology and API provides support for structured metadata and controlled vocabularies wherever possible, but they do not enforce their use.

Extensibility: Because of the Web's flexible structure, future versions of the specification may contain additional properties in the core vocabulary (and its representations) and mappings to additional metadata formats.

Interoperability: Syntactic and semantic interoperability is ensured by the defined semantics of the set of core properties and the mapping tables to the metadata formats.

Granularity: The ontology and API can be used independently, depending on the actual application domain. Furthermore, conformance to the specifications is possible on different levels of strictness.

Using these requirements as a foundation, we can provide more details on the ontology and API.

Ontology for Media Resources 1.0

The set of core properties that constitutes this ontology is based on a list of the most commonly used annotation properties from current multimedia metadata schemas. The properties were derived from the *Multimedia Vocabularies on the Semantic Web* report of the Multimedia Semantics Incubator Group (see www.w3.org/2005/Incubator/mmsem/XGR-vocabularies-20070724). We also compiled a list of use cases from a public call. Over the course of the group's work, we have mapped these core properties to more than 15 metadata and container formats from a range of communities.

The set of core properties defined in the Ontology for Media Resources 1.0 consists of 20 descriptive metadata properties (such as identifiers, language, contributors, creation date, genre, and ratings) and eight technical metadata properties (such as frame size, duration, and format). The descriptive properties are media agnostic and apply to descriptions of multimedia works (such as movies) that are not specific to an instantiation (an AVI file, for example). The technical properties, bound to certain media types, are essential only when describing a particular instantiation of the content. Following the requirements of a core vocabulary, all properties are defined with explicit semantics to clarify and disambiguate their definitions in the context of a media resource description.

Whenever these properties exist in other standards, as we outlined earlier, the Ontology

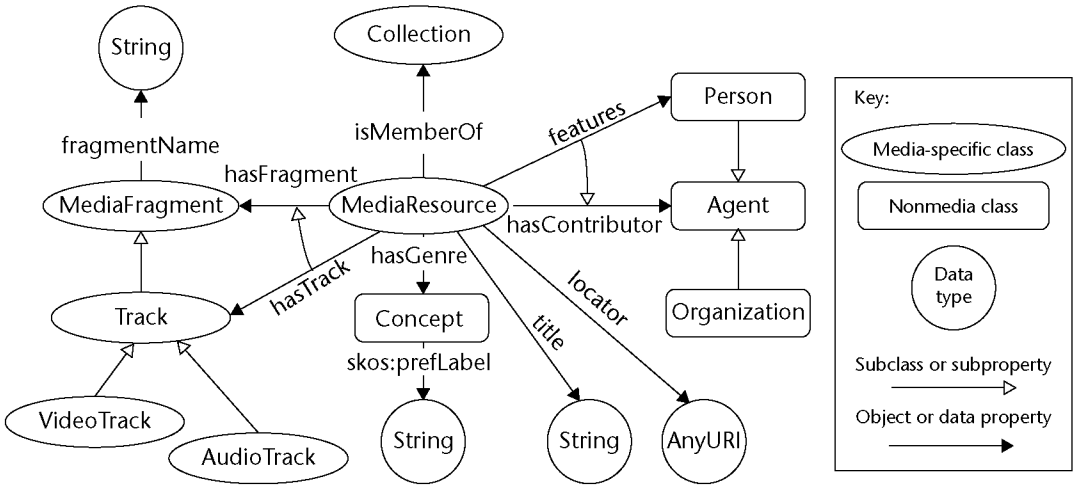


Figure 1. Class model excerpt of the Ontology for Media Resources 1.0. The class model defines a set of media- and nonmedia-specific classes.

Table 1. Supported metadata and container formats.

Format type	Supported formats
Multimedia metadata format	CableLabs 1.1, DIG, Dublin Core, EBUCore, EXIF, ID3, IPTC NewsML-G2, LOM, MediaRDF, MediaRSS, METS, MPEG-7, OGG, QuickTime, SMTPD, TV-Anytime, TXFeed, XMP, and YouTube
Multimedia container formats	3GP, Flash/FLV, Flash/F4V, MPEG-4 (MP4), MOV (Quicktime), OGG, and WebM

for Media Resources 1.0 specification document explicitly defines how they are related. Furthermore, the ontology can be used with different layers of conformance. If an extension of the basic property semantics is necessary, optional subtypes can be used to further qualify many of the descriptive properties—for example, to define a specific kind of contributor.

Another aim of MAWG was to develop a Semantic Web-compatible implementation of the Ontology for Media Resources 1.0 based on the W3C recommendations Resource Description Framework (RDF) and the Web Ontology Language (OWL). (The latest version of the RDF/OWL implementation is available at http://www.w3.org/ns/ma-ont.) The ontology's domain analysis is based on the list of properties and the mappings defined by the group. The actual conceptualization and implementation of the Semantic Web compatible ontology is based on them. The identified class model, which is based on the EBU Class Conceptual Data Model (CCDM, http://tech.ebu.ch/classmodel) for distribution, defines a set of media- and nonmedia-specific classes.

Figure 1 illustrates an excerpt of the Ontology for Media Resources 1.0 class model. (See earlier work for a detailed description of the conceptual model, implementation, and use of the Ontology for Media Resources 1.0.[13])

The set of properties modeled in the ontology corresponds to existing metadata formats currently describing media resources published on the Web. These correspondences have been defined in the form of mappings to provide an interoperable set of metadata, thereby enabling different applications to share and reuse metadata. The selected metadata formats were those we deemed to be the most commonly adopted. Specifically, we selected the 19 media metadata formats and seven media container formats that are listed in Table 1. This list of formats is not closed, nor does it pretend to be exhaustive. A future version of the Ontology for Media Resources 1.0 may include additional mappings if a new need or use case is identified. In this light, current work focuses on creating additional mappings for Schema.org and Atom. (See www.w3.org/2008/WebVideo/Annotations/drafts/ontology10/additional-mappings.html for more details.)

The mappings defined between the Ontology for Media Resources 1.0 properties and those of the selected metadata formats are available in metadata-format-specific mapping

tables (see www.w3.org/TR/2012/REC-mediaont-10-20120209/#property-mapping-table). The mappings (defined as one-way mappings from the source format to the core properties) may have different semantic relations. To express this, we use the following subset of SKOS semantic relations: exactMatch, broadMatch, narrowMatch, and relatedMatch. (See www.w3.org/TR/skos-reference/#semantic-relations for more details.)

API for Media Resources 1.0

The API for Media Resources 1.0 enables interoperable access to metadata information related to media resources on the Web, with the defined core vocabulary as a recommended best practice. The MAWG discussed various design considerations, which led to the specification of global interfaces with specific parameters. This implies a minimal number of exposed interfaces ensuring a broad adoption and fewer security leaks. Furthermore, it reduces implementation work when designing applications or integrating the API into legacy systems.

The API can be used in two modes of operation: asynchronous and synchronous. For this API, the asynchronous mode is considered the default, where calls return without waiting for the request to finish its execution. A call-back function is provided to be invoked when the request terminates. On the other hand, synchronous calls wait for the request to terminate and directly return the result. Figure 2 illustrates the two scenarios for the API's use. In the first scenario, the API is encapsulated in a user agent, whereas in the second, it is implemented as a Web service.

This API has two main parts: interfaces to access media resources and a set of core properties describing the information in an interoperable way along with their JavaScript Object Notation (JSON) serialization. The API is defined using WebIDL (www.w3.org/TR/WebIDL), which is an Interface Description Language (IDL) variant explicitly covering programming languages commonly used on the Web (such as ECMAScript). Following the ontology design, every property can carry its metadata information in either an unstructured (single value) or structured way.

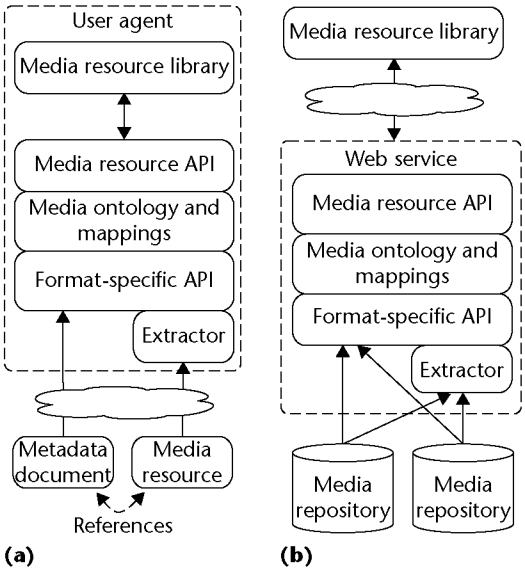


Figure 2. Design considerations of the API for Media Resources. In these two scenarios, the API is either encapsulated in a user agent or implemented as a Web service.

Projects and Showcases

As we mentioned earlier, the MAWG's goal is unified access to media resources on the Web. Recent works in the social media area heavily utilize media resources in their workflows, aiming to improve interactive communication between peers. This domain consists of three main parties: content providers, portal services, and consumers. Canonical processes define the connection between these parties. Here, the MAWG standardization effort can be directly injected into those processes to improve interoperability with respect to data exchange.

In this light, Table 2 gives an overview of current scientific projects implementing

Table 2. Projects utilizing the MAWG specifications.

Project name	Type	Ontology	Mappings	API
PrestoPRIME Semantic Converter[14] (see Figure 3)	Metadata mapping service	X	X	–
Multimedia Metadata Ontology[15]	Metadata mapping service	X	–	–
Linked Media Framework[16] (see Figure 4)	Portal service	X	X	X
NinSuna[17] (Figure 5)	Portal service	X	X	–
EventMedia[18] Live (Figure 6)	Portal service	X	–	–

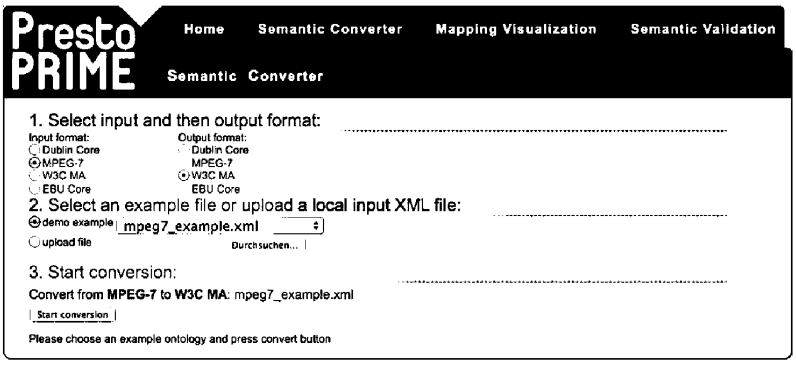


Figure 3. PrestoPRIME Semantic Converter online demonstrator. This service uses the Ontology for Media Resources as an interoperable target format.

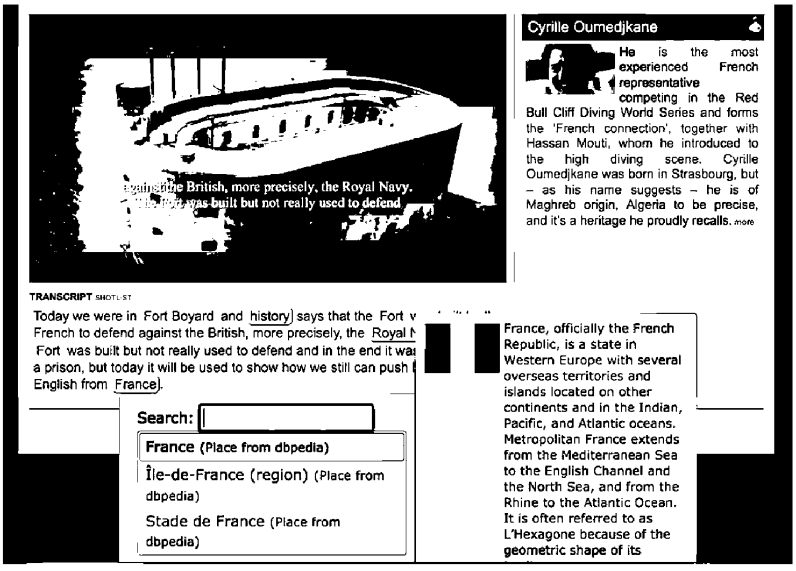
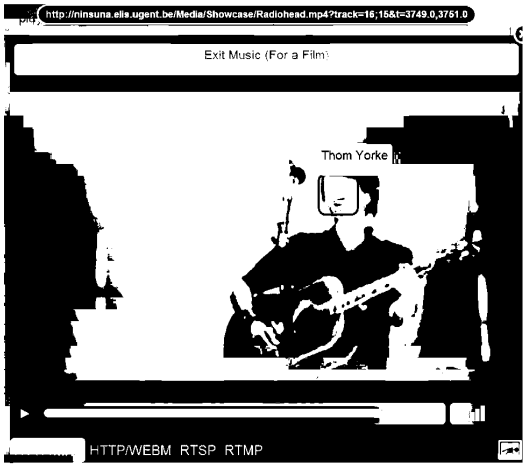


Figure 4. Linked Media Framework. A media interlinking module uses the Ontology for Media Resources and the W3C URI for Media Fragments to enable integration with heterogeneous data sources on the Web.

Figure 5. NinSuna Radiohead showcase. The Ontology for Media Resources is a driver for interoperability, and it allows the integration of other data sources within NinSuna.



MAWG specifications, and Figures 3 through 6 depict their online demonstrators. The table shows that diverse projects can integrate the specifications with different layers of conformance.

The PrestoPRIME Semantic Converter is an automated metadata mapping service for audio-visual metadata in the archival domain.[14] It uses the Ontology for Media Resources as an interoperable target format (see Figure 3). The service supports a number of metadata standards, and proprietary formats from archive and broadcast organizations can be added. It is becoming increasingly important to make archive content accessible on the Web together with its metadata, making the Ontology for Media Resources a relevant target format for publication and interoperability with linked data.

The Multimedia Metadata Ontology (M3O) is a comprehensive model for representing multimedia metadata.[15] It is based on the foundational ontology DOLCE+DnS Ultralight and several ontology design patterns that have been aligned with the Ontology for Media Resources. M3O serves as a generic modeling framework for integrating existing metadata models and metadata standards rather than replacing them, also providing support for the Ontology for Media Resources.

The Linked Media Framework (LMF) is an easy-to-setup server application that bundles central Semantic Web technologies.[16] Its goal is to offer publishing legacy data as linked data, building semantic search applications and enabling information extraction. A number of additional modules are provided in the LMF, among them the media interlinking module, which uses the Ontology for Media Resources and the W3C URI for Media Fragments to enable integration with heterogeneous data sources on the Web (see Figure 4).

NinSuna is a metadata-driven media adaptation and delivery framework.[17] This framework uses novel media support in HTML5 and supports fragment-based access conforming to the W3C URI for Media Fragments specification. Metadata for the media items is published in RDF conforming to the Ontology for Media Resources, providing powerful time- and region-based annotation capabilities in combination with fragment identifiers (see Figure 5). The use of the Ontology for

Media Resources is a driver for interoperability and allows the integration of other data sources within NinSuna.

EventMedia aggregates a large dataset that consists of event descriptions (from the public event directories last.fm, eventful, and upcoming) together with media descriptions associated with these events and interlinked with the larger Linked Open Data Cloud.[18] A Web-based environment lets users explore and select events and view associated media. The Ontology for Media Resources has been used for representing the metadata of these media and has enabled interlinking with the Linked Open Data cloud (see Figure 6).

All these projects use the core vocabulary to describe a media resource's information in a unified way. In addition to supporting the core vocabulary, the PrestoPRIME Semantic Converter, LMF, and NinSuna also implement the mappings defined by the MAWG—using Extensible Stylesheet Language Transformations (XSLT), for example. These applications thus bridge the interoperability gap by providing multimedia metadata published in different source formats in a unified way. LMF further implements the API to enable a unified retrieval over the heterogeneous landscape of metadata formats available in the Linked Data Cloud.

In addition to these projects, the following two open source showcases are a starting point for future implementations focusing on the API for Media Resources 1.0. The first deals with a gallery showing images as well as its metadata information. (Both implementations are available online at http://mawg.joanneum.at.) Here, the API is implemented as a Web service using the synchronous operation mode. The second showcase utilizes the API in a browser extension following the asynchronous operation mode. The application lets a user generate a video playlist, where videos and the corresponding metadata information from different platforms can be arranged in a unified way. These implementations serve as a validation for the API specification, provide useful feedback for the specification, and confirm its applicability. In addition, the code from these implementations provides a convenient starting point for developers interested in implementing the API.

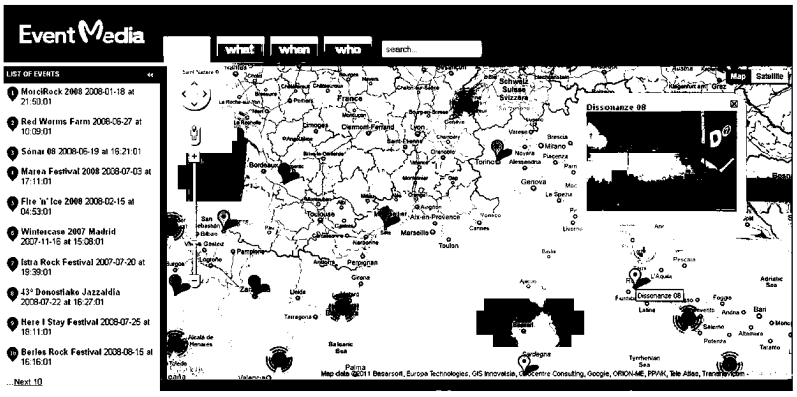


Figure 6. EventMedia. The Ontology for Media Resources allows EventMedia to represent the metadata of public event media and enables interlinking with the Linked Open Data Cloud.

Conclusion

The bottom line is that both specifications have been effectively implemented by several institutes and justified by standardized test suites available on the working group page. The Ontology for Media Resources 1.0 is an official W3C recommendation, and the API is currently a proposed recommendation and is expected to become a W3C recommendation by the middle of 2013. **MM**

Acknowledgments

We thank the W3C MAWG Chairs Soohong Daniel Park and Joakim Söderberg, the W3C team contact Thierry Michel, and all the MAWG participants for their contributions as well as Yiannis Kompatsiaris for his valuable feedback on this article. This work has been partially supported by the German Federal Ministry of Economics and Technology under the THESEUS Program and by the European Union's 7th Framework Programme (FP7/2007-2013) under grant agreement 231161, PrestoPRIME.

References

1. F. Nack, "All Content Counts: The Future in Digital Media Computing Is Meta," *IEEE Multimedia,* vol. 7, no. 3, 2000, pp. 10–13.
2. H. Kosch et al., "The Life Cycle of Multimedia Metadata," *IEEE Multimedia,* vol. 12, no. 1, 2005, pp. 80–86.
3. J.R. Smith and P. Schirling, "Metadata Standards Roundup," *IEEE Multimedia,* vol. 13, no. 2, 2006, pp. 84–88.
4. L. Hardman et al., "Canonical Processes of Semantically Annotated Media Production," *Multimedia Systems,* vol. 14, no. 6, 2008, pp. 327–340.
5. J.R. Smith, "The Search for Interoperability," *IEEE Multimedia,* vol.15, no. 3, 2008, pp. 84–87.
6. M.C. Suàrez-Figueroa, G.A. Atemezing, and O. Corcho, "The Landscape of Multimedia

Ontologies in the Last Decade," *Multimedia Tools and Applications,* vol. 62, no. 2, 2011, pp. 1–23.
7. R. Troncy et al., "MPEG-7 Based Multimedia Ontologies: Interoperability Support or Interoperability Issue?" *Proc. 1st Workshop on Multimedia Annotation and Retrieval Enabled by Shared Ontologies,* 2007.
8. S. Dasiopoulou et al., "Enquiring MPEG-7 Based Multimedia Ontologies," *Multimedia Tools Applications,* vol. 46, nos. 2–3, 2010, pp. 331–370.
9. B. Haslhofer and W. Klas, "A Survey of Techniques for Achieving Metadata Interoperability," *ACM Computing Surveys,* vol. 42, Mar. 2010, pp. 7:1–7:37.
10. S.D. Park et al., "Video in the Web: Technical Challenges and Standardization," *IEEE Multimedia,* vol. 17, no. 4, 2010, pp. 90–93.
11. W. Lee et al., *Ontology for Media Resources 1.0,* World Wide Web Consortium (W3C) recommendation, Feb. 2012; www.w3.org/TR/2012/REC-mediaont-10-20120209.
12. F. Stegmaier et al., "API for Media Resources 1.0," World Wide Web Consortium (W3C) recommendation, Nov. 2011; www.w3.org/TR/2011/WD-mediaont-api-1.0-20111122.
13. J.-P. Evain and T. Bürger, "Semantic Web, Linked Data and Broadcasting — More in Common than You'd Think!" *EBU Tech. Rev.,* vol. 2011, no. Q1, 2011, pp. 1–13.
14. M. Höffernig et al., "Mapping Audio-Visual Metadata Formats Using Formal Semantics," *Proc. 5th Int'l Conf. Semantic and Digital Media Technology* (SAMT), LNC S 6725, Springer, 2011, pp. 80–94.
15. A. Scherp, D. Eißing, and C. Saathoff, "A Method for Integrating Multimedia Metadata Standards and Metadata Formats with the Multimedia Metadata Ontology," *Int'l J. Semantic Computing,* vol. 6, no. 1, 2012, pp. 25–49.
16. T. Kurz, S. Schaffert, and T. Bürger, "LMF: A Framework for Linked Media," *Proc. Workshop on Multimedia on the Web (MMWEB),* IEEE CS, 2011, pp. 16–20.
17. W. Van Lancker et al., "Harmonizing Media Annotations and Media Fragments," *Proc. Workshop on Multimedia on the Web (MMWEB),* IEEE CS, 2011, pp. 1–4.
18. R. Troncy, B. Malocha, and A.T.S. Fialho, "Linking Events with Media," *Proc. 6th Int'l Conf. Semantic Systems* (I-SEMANTICS), ACM, 2010, article no. 42.