

## View other issues

### Contents: ERCIM News 72

## Creating, Organising and Publishing Media on the Web

by Raphaël Troncy

**Creating, organising and publishing findable media on the Web raises outstanding problems. While most applications that process multimedia assets make use of some form of metadata to describe the multimedia content, they are often based on diverse metadata standards. In the W3C Multimedia Semantics Incubator Group (XG), we have demonstrated, for various use cases, the added value of combining several of these into a single application. We have also shown how semantic Web technology can help in making these standards interoperable.**

The W3C Multimedia Semantics XG has been chartered to show how metadata interoperability can be achieved by using semantic Web technology to integrate existing multimedia metadata standards. Between May 2006 and August 2007, 34 participants from fifteen different W3C organisations and three invited experts have produced several documents: these describe relevant multimedia vocabularies for developers of semantic Web applications, and the use of RDF and OWL to create, store, exchange and process information about image, video and audio content. The Semantic Media Interfaces group at CWI has sponsored and co-chaired this W3C activity.

The main result is a Multimedia Annotation Interoperability Framework that presents several key use cases and applications showing the need for combining multiple multimedia metadata formats, often XML-based. We use semantic Web languages to explicitly represent the semantics of these standards and show that they can be better integrated on the Web. We describe two of the use cases here.

### Managing Personal Digital Photos

The advent of digital cameras and camera phones has made it easier than ever before to capture photos, and many people have collections of thousands of digital photos from friends and family, or taken during vacations, while travelling and or parties. Typically, these photos are stored on personal computer hard drives in a simple directory structure without any metadata. The number of tools, either for desktop machines or Web-based, that perform automatic and/or manual annotation of content is increasing. For example, a large number of personal photo management tools extract information from the so-called EXIF header and add this information to the photo description. Popular Web 2.0 applications, such as Flickr, allow users to upload, tag and share their photos within a Web community, while Riya provides specific services such as face detection and face recognition of personal photo collections.

What remains difficult however, is finding, sharing, and reusing photo collections across the borders of tools and Web applications that come with different capabilities. The metadata can be represented in many different standards ranging from EXIF (for the camera description), MPEG-7 (for the low-level characteristics of the photo) to XMP and DIG35 (for the semantics of what is depicted). To address this problem, we have designed OWL (Web Ontology Language) ontologies to formalize the semantics of these standards, and we provide converters from these formats to RDF in order to integrate the metadata. A typical use case is a Web application that can generate digital photo albums on the fly based on both contextual and semantic descriptions of the photos, and aesthetic and visual information tailored to the user.

### Faceting Music

Many users own digital music files that they have listened to only once, if ever. Artist, title, album and genre metadata represented by ID3 tags stored within the audio files are generally used by applications to help music consumers manage and find the music they like. Recommendations from social networks, such as lastFM, combined with individual FOAF (Friend of a Friend) profiles comprise additional metadata that can help users to purchase missing items from their collections. Finally, acoustic metadata, a broad term which encompasses low-level features such as tempo, rhythm and tonality (key), are MPEG-7 descriptors that can be automatically extracted from audio files, and which can then be used to generate personalized playlists.

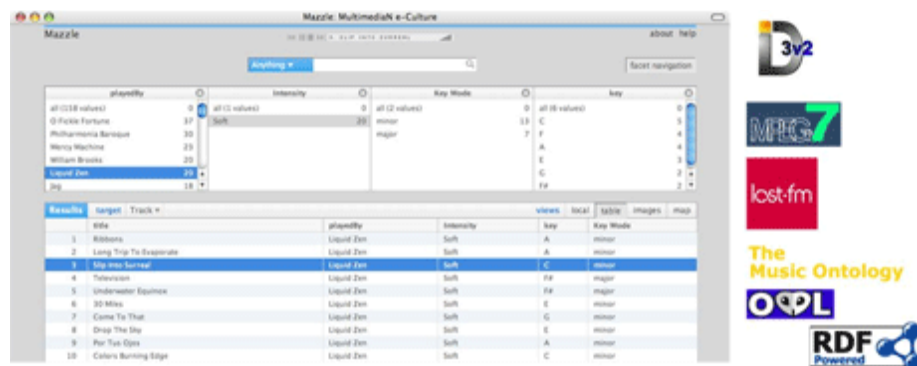


Figure 1: Muzzle provides a faceted browser interface on your personal music collection and can automatically build personalized playlists.

Using the Music Ontology, a converter generates RDF descriptions of music tracks from the ID3 tags. The descriptions are enhanced with low-level descriptors extracted from the signal and can be linked to recommendations from social community Web sites. The user can then navigate a personal music collection using an iTunes-like faceted interface that aggregates all this metadata (see Figure 1).

### COMM: From MPEG-7 to the Semantic Web

MPEG-7 is the de facto standard for multimedia annotation that appears in every use case developed by the Multimedia Semantics XG. Media content descriptions often explain how a media object is composed of its parts and what the various parts represent. Media context descriptions need to link to some background knowledge and to the creation and use of the media object. To make MPEG-7 interoperable with semantic Web technology, we have designed COMM, a core ontology for multimedia that has been built by reengineering the ISO standard, and using DOLCE (Descriptive Ontology for Linguistic and Cognitive Engineering) as its underlying foundational ontology to support conceptual clarity and soundness as well as extensibility towards new annotation requirements.

Current semantic Web technology is sufficiently generic to support annotation of a wide variety of Web resources, including image, audio and video resources. The Multimedia Semantics XG has developed tools and provided use cases showing how the interoperability of heterogeneous metadata can be achieved. Still, many things need to be improved! The results from the Multimedia Semantics XG and the report from the W3C 'Video on the Web' workshop will provide valuable input to identifying the next steps towards making multimedia content a truly integrated part of the Web.

### Links:

<http://www.w3.org/2005/Incubator/mmsem/>  
<http://multimedia.semanticweb.org/COMM/>  
<http://www.w3.org/2007/08/video/>

### Please contact:

Raphaël Troncy, CWI, The Netherlands  
 Tel: +31 20 592 4093  
 E-mail: Raphael.Troncy@cwi.nl