Deploying Multimedia Metadata
on the Semantic Web

Michael Hausenblas, Werner Bailer, Tobias Bürger, and Raphaël Troncy

Abstract—The effective and efficient deployment of multimedia metadata along with the content is a critical shortcoming regarding multimedia applications on the Semantic Web. A large set of multimedia metadata formats have been used for years in diverse applications. The deployment of these formats in the context of the Semantic Web is, however, problematic since a framework allowing agents to directly access and process existing multimedia metadata descriptions is currently missing. In this paper, we propose a non-disruptive way of deploying multimedia metadata on the Web using RDFa along with a lightweight formalisation vocabulary.

I. INTRODUCTION

Many multimedia metadata (M3) formats, such as ID3, Exif or MPEG-7 are available to describe what a multimedia asset is about, who has produced it, how it can be decomposed, etc. [1]. For professional content found in archives and digital libraries, a range of in-house or standardised M3 formats is used. Similar issues arise with the dissemination of user generated content found at social media sites such as Flickr (http://www.flickr.com), YouTube (http://www.youtube.com), or Facebook (http://www.facebook.com).

To enable the deployment of existing multimedia metadata formats on the Semantic Web, we advocate the use of the RDF data model [2]. A large number of existing M3 formats have been used for years in diverse applications. Instead of discarding these multimedia metadata standards, we propose a solution that allows hooking existing multimedia metadata formats into the Semantic Web: *RDFa-deployed Multimedia Metadata (ramm.x)*.

ramm.x describes media assets published on the Web by linking existing descriptions represented in a M3 format and referencing services capable of transforming parts of (or whole) descriptions into RDF. The step of RDFising is called *formalisation*. It is typically carried out using a multimedia ontology [1, Sec. 4]).

We propose to use RDFa [3] to actually deliver the metadata along with the content being served. RDFa is a serialisation syntax for the RDF data model intended to be used in (X)HTML environments. It defines how an RDF graph is embedded in an (X)HTML page using a set of defined attributes such as `@about`, `@rel`, `@instanceof`, etc. Hence, the

M. Hausenblas and W. Bailer are with the Institute of Information Systems & Information Management, JOANNEUM RESEARCH Forschungsges. mbH, Steyrergasse 17, 8010 Graz, Austria, {firstname.lastname}@joanneum.at

T. Bürger is with Digital Enterprise Research Institute (DERI), Techniker-strasse 21a, 6020 Innsbruck, Austria, tobias.buerger@deri.at

R. Troncy is with CWI Amsterdam, P.O. Box 94079, 1090 GB Amsterdam, The Netherlands, raphael.troncy@cwi.nl
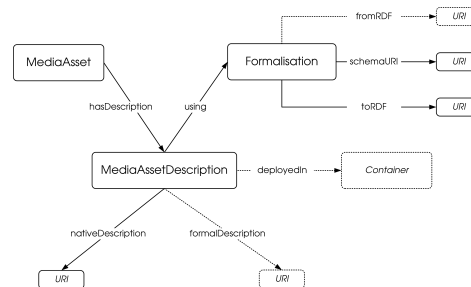

Fig. 1. Overview on ramm.x core vocabulary.

basic idea of ramm.x is to define a light-weight deployment description vocabulary allowing—in conjunction with RDFa—a Semantic Web agent to determine the formalisation steps in order to process the native M3 format.

II. OVERVIEW ON RAMM.X

Potential applications of ramm.x have the following properties: (i) media assets are published on the Web, (ii) they are published with metadata describing them, (iii) their metadata is not just free text or natively represented using an RDF-based ontology and (iv) there is an added value from making the multimedia metadata available to a Semantic Web agent. A typical use case for ramm.x is the description of videos. Video applications, especially if using complex descriptions of its scene structures, greatly benefit from ramm.x. In a typical Web-based scenario, the whole clip is described with some tags. However, by accessing an MPEG-7 [4] document describing the video structure and the content of the segments, a ramm.x client could link shots of a video to related information available on the Web, such as resources about people or places appearing in this segment.

A possible ramm.x application domain is the description of cultural heritage assets: the applicability and benefits for the description of newspaper archives have already been shown in [5].

A. Core Vocabulary

The ramm.x core vocabulary[1] defines the following concepts (Fig. 1): A **media asset** published on the Web is an information resource[2] identified by its URL[3]. **Media asset descriptions** (MAD) are the pivot element for bridging the gap between a
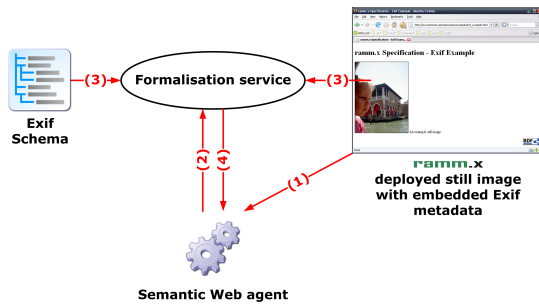
---

[1]See also http://sw.joanneum.at/rammx/spec/
[2]http://www.w3.org/2001/tag/doc/httpRange-14/HttpRange-14.html
[3]URI declaration vs. usage, http://dbooth.org/2007/uri-decl/

Fig. 2. Demonstration of ramm.x processing.

```
1  <div about="#exif_formal"
       instanceof="ramm:Formalisation">
2   <span rel="ramm:schemaURI"
        href="http://www.kanzaki.com/ns/exif"/>
3   <span rel="ramm:toRDF"
        href="http://www.kanzaki.com/test/exif2rdf"/>
4  </div>
5  <div about="#sample_mad"
       instanceof="ramm:MediaAssetDescription">
6   <span about="#exif_example.jpg"
        instanceof="ramm:MediaAsset">
7    <img src="exif_example.jpg" />
8    <span rel="ramm:hasDescription"
         href="#sample_mad"/>
9   </span>
10  <span rel="ramm:nativeDescription"
        href="#exif_example.jpg"/>
11  <span rel="ramm:using" href="#exif_formal"/>
12 </div>
```

Listing 1. A ramm.x deployed still image.

multimedia metadata format in its native representation and the formalised version usable on the Semantic Web. Each media asset description points exactly to one native description, and uses exactly one formalisation. Optionally, there might be a formalised version of the description available as well. A **formalisation** is an explicit representation of the transformation process regarding a specific schema (in RDF-S, OWL, etc.). It refers to URLs pointing to converter services to and from an RDF-based representation, along with the respective schema. A **container** is an optional logical unit used to bundle media asset descriptions. It is therefore possible to assign properties to a set of media asset descriptions using a container element, for example, a copyright note using `cc:license` attached to the container.

B. Processing

A Semantic Web agent needs to process the container document in order to use the ramm.x deployed metadata:

1) The Semantic Web agent fetches the ramm.x container (e.g. an XHTML document) with the embedded metadata and extracts the RDF using an RDFa extractor[4].
2) The formalisation service is detected and invoked, e.g. by lookup of a Web service, preparation of a SOAP message, etc.
3) Using the `toRDF` property of the `Formalisation` along with the defined schema (from `schemaURI`), the original multimedia metadata format—denoted by the value of `nativeDescription`—is converted to its RDF representation by the formalisation service.
4) The resulting RDF graph is transferred back to the Semantic Web agent and can be further used in a query, or an inference process.

III. DEMONSTRATION

A simple demonstration of ramm.x is shown in the following description of a still image containing embedded Exif metadata[5]. In listing 1 the ramm.x deployment is depicted: First, a media asset is described identified by its URL (line 6). The media asset has a description `#sample_mad` (line 5) referring to the native Exif metadata embedded in the still

image. Using the formalisation `#exif_formal` (line 1), the full formal description can be obtained. The formal description conforms to a schema, and can be generated using a formalisation service. Finally, the processing of the ramm.x description is depicted in Fig. 2.

IV. CONCLUSION AND OUTLOOK

Based on the ramm.x use cases[6], we have identified several issues that will be addressed by future extensions[7]: (i) collect formalisations of different multimedia metadata vocabularies together with the services capable of doing the formalisation, (ii) deal with the scalability of the deployment of large media asset descriptions such as a detailed annotation of a one hour documentary (a "streaming mode" allowing to access parts of the description), and (iii) trace orphaned media assets: when a media asset is taken out of its original context, one would certainly expect to not lose the associated metadata; this can be achieved by the use of watermarking techniques carrying a pointer to the original container URI.

REFERENCES

[1] M. Hausenblas eds., "Multimedia Vocabularies on the Semantic Web," W3C Multimedia Semantics Incubator Group Report, 2007.
[2] G. Klyne, J. Carroll, and B. McBride, "Resource Description Framework (RDF): Concepts and Abstract Syntax," W3C Recommendation, 2004.
[3] B. Adida and M. Birbeck, "RDFa Primer 1.0," W3C Semantic Web Deployment Working Group, Working Draft, 2007.
[4] MPEG-7, "Multimedia Content Description Interface," Standard ISO/IEC 15938, 2001.
[5] M. Hausenblas, W. Bailer, and H. Mayer, "Deploying Multimedia Metadata in Cultural Heritage on the Semantic Web," in *International Workshop on Cultural Heritage on the Semantic Web*, Busan, South Korea, 2007.

[4]http://esw.w3.org/topic/RDFa#Implementations

[5]The complete demonstration is available via the ramm.x specification at http://sw.joanneum.at/rammx/spec/example/exif_example.html

[6]http://sw.joanneum.at/rammx/usecases/

[7]http://sw.joanneum.at/rammx/extensions/