DESIDOC Journal of Library & Information Technology, Vol. 31, No. 4, July 2011, pp. 247-252

Advanced (X) HTML5 Metadata and Semantics for Web 3.0 Videos

Leslie F. Sikos

IT Professional, Adelaide, Australia
E-mail: lesliesikos@gmail.com

ABSTRACT

Searching for videos online is inefficient if the search relies exclusively on relevant keywords in the text surrounding the object. HTML5 introduced advanced semantic annotation in the form of a new metadata format called microdata. Along with the video element which provides semantics to embedded video contents, media metadata can also be used to add additional features to be searched, archived, or automatically processed. Certain metadata formats provide a high level of flexibility adequate to combine them with other vocabularies, making it possible to write truly sophisticated, machine-readable metadata for video components of the Social Semantic Web.

Keywords: Machine-readable metadata, video embedding, microformats, microdata, Dublin Core, RDFa, HTML5, social Semantic Web

1. INTRODUCTION

For years, publishing videos online was feasible through general object embedding only. Currently, there is still no standardised solution for publishing videos on the Web, partly due to the variety of video and audio codecs, and the varying browser support. Video metadata can be embedded either in the markup or in the file itself. Without descriptive, structured metadata video searches are insufficient, unreliable, or even unfeasible.

The first video metadata applications date-back before the Web 2.0 era when Dublin Core metadata was applied on MPEG-7 videos to support hierarchical video descriptions[1], and later for providing conceptual navigation[2]. Although the need for video sharing has increased enormously with the introduction of Semantic Web and especially the Social Semantic Web, the potential of Semantic Web technologies in video sharing has been far from exploited yet although some efforts have already been made in this field.

Gómez-Romero, *et al*. used ontologies to build a framework to construct a symbolic model of video scenes by integrating tracking data and contextual information[3]. Brostow, *et al*. presented the first collection of videos with object class semantic labels, complete with metadata[4]. Their database associates each pixel with one of 32 semantic classes which they demonstrated on three distinct fields. Gayo, *et al*. introduced video metadata for both the producers and the users[5]. The first type of metadata associates multimedia content with domain-specific ontologies while the second provides video tagging. Their search engine combines both kinds of metadata to locate the desired content and provides browsing capabilities as well. Kucuk, *et al*. proposed an annotation and retrieval system to automate semantic video indexing[6]. Along with the introduction of HTML5, new opportunities appeared to provide rich semantics as well as accessibility in the markup for videos. This paper enumerates the challenges as well as limitations of the field together with methods for standard-compliant, semantic video embedding, and metadata mixing.

2. SEMANTIC VIDEO EMBEDDING

If an online video is not meant to be part of an interactive multimedia presentation, that is best provided by SMIL along with its native metadata[7], the resource is described by (X)HTML markup. In contrast to the complexity of video embedding with the embed element in HTML versions up to 4.01, or the object element in XHTML, (X)HTML5 introduced the video element that provides semantic meaning to video contents and full control over the embedded file[8]. Features such as height or width can be added optionally. An image representing a frame from the video can also be defined in case the video cannot be rendered. Alternate content can also be given in the form

```
<video src="sample.ogv" width="320" height="240"
poster="sample.jpg">
    <p>Download the <a href="video.ogv">sample video</
      a> (OGV, 5.34 MB)</p>
</video>
```

Video controls can be shown or hidden in the browser by the controls attribute on the video element:

```
<video src="xyz.mov" controls="controls">
</video>
```

Since the video codec support is different in each browser, the same video can be provided in various formats, avoiding the need to download videos that cannot be played on the system, e.g.,

```
<video>
    <source src="video.mp4" type="video/mp4">
    <source src="video.ogv" type="video/ogg">
    <p>Download the <a href="video.ogv">sample video</
      a> (OGV)</p>
</video>
```

The type attribute declares the Internet Media Type (MIME type) of the video file. These specifications are described in IETF/ISOC Request for Comments (RFC) documents and are standardised by IANA registration[9]. The most common video Media Types are summarised in Table 1.

Table 1. Common MIME types for video

Media type	Specification	Description
video/mpeg	RFC2045[10], RFC 2046[11]	MPEG-1 video
video/mp4	RFC 4337[12]	MP4 video
video/ogg	RFC 5334[13]	Ogg theora or other video
video/quicktime	IANA registered[14]	QuickTime video
video/x-ms-wmv	MS KB 288102[15]	Windows media video

The problem is that MIME types do not reflect the codecs of videos stored in container formats (e.g., H.264 in MPEG-4). They can be provided by the codecs parameter:

```
<video controls="controls">
    <source src="video.mp4" type='video/mp4;
      codecs="avc1.42E01E, mp4a.40.2" '>
    <source src="video.ogv" type='video/ogg;
      codecs="theora, vorbis" '>
    <p>Download the <a href="video.ogv">sample video</
      a></p>
</video>
```

The video element of (X)HTML5 provides playback support detection, including the canPlayType() method on the media element[16] or the onerror event listener. Certain browsers cannot stream the video or automatically download the whole video file even if playback has not been started yet. One of the exceptions is Firefox 3.6+ which only downloads a fragment necessary to determine duration and render a frame from the video. This behaviour can be overridden by the preload attribute. The attribute value none forces the browser to avoid downloading. The metadata attribute value hints that enough data should be downloaded only to show a frame and determine duration. The value auto downloads the whole video if possible. The effect of preload="none" can be simulated in browsers that do not support it by omitting the src attribute and source elements that are provided only if the user clicks a button:

```
<video controls="controls">
   Video not supported
</video>
<input type="button" value="Load video"
  onclick="document.getElementsByTagName('video')[0].src
  = 'video.mp4';">
```

Additionally, customised controls can be added to the video embedding since the DOM API for video in (X)HTML5 supports several events that can be implemented through JavaScript. For example:

```
<script>
 var video = document.getElementsByTagName
 ('video')[0];
</script>
<input type="button" value="Play"
 onclick="video.play()">
<input type="button" value="Pause"
 onclick="video.pause()">
```

Currently, the src attribute value of the video tag should be a physical file which makes it impossible to embed videos from YouTube.

(X)HTML5 videos can be dynamically drawn on a canvas with JavaScript using the drawImage method:

```
<video src="video.mp4" controls="controls">
   Video not supported
</video>
  <canvas id="canvas">
   Canvas not supported
</canvas>
<script>
var ctx = document.getElementById('canvas').
getContext('2d');
var video =
document.getElementsByTagName('video')[0];
video.onloadeddata = function(e) {
ctx.canvas.width = video.videoWidth;
ctx.canvas.height = video.videoHeight;
ctx.drawImage(video, 0, 0);
  }
</script>
```

3. MACHINE-READABLE METADATA

Certain metadata formats such as microformats and RDFa apply and reuse features of existing technologies (e.g., the rel attribute)[17] while others such as Creative Commons and Dublin Core introduce new annotations, typically through the namespace mechanism.

3.1 Metadata Formats Reusing Attributes

The Web document attributes used by RDFa are ideal for metadata on generic objects as well as videos[18]. A common feature of video sharing portals is the review option which can be implemented with metadata support by using the hReview microformat (class=”hreview”).

Some or all rights of video authors are often reserved. Many licenses associated with videos are sophisticated and users cannot be expected to know them. The rel=”license” microformat can be added to hyperlinks that point to the description of the license. In the case of the Creative Commons Attribution-ShareAlike license, for example, rel=”license” should be

```
<video src=”lesson1.ogv” rel=”license” href=”http://
creativecommons.org/licenses/by-sa/3.0/” />
```

The value of the *href* attribute provides the associated URI of the resource where the license is described.

3.2 Video Licensing with HTML5 Microdata

The concept of microdata has been introduced in HTML5 for labelling content in order to describe a specific type of information, such as licensing information to eliminate copyright issues and to contribute to advanced Web searches[19]. For example:

```
<footer>
  <p id=”licenses”>All videos are licensed under the <a
    itemprop=”license”
    href=”http://creativecommons.org/
    licenses/by-sa/3.0/”>Creative Commons
    Attribution Share Alike license</a>.</p>
</footer>
```

3.3 Metadata from External Vocabularies

Being a general Web resource description and modelling language, Resource Description Framework (RDF) can be used for describing any kind of resources that can be identified by a URI[20], including video files.

RDFa (RDF in attributes) provides the option to embed rich metadata within certain attributes of Web documents[21]. The set of attributes to be used for this purpose are about, src, rel, rev, href, resource, property, content, datatype, and typeof, all of which are supported in (X)HTML5. RDFa is indexed by Google and Yahoo and supported by the highest traffic video sharing website YouTube.

The Dublin Core vocabulary has several metadata that can be used for video contents such as MIME type, date and time, associated URI, country, and language[22].

Video licenses can be efficiently denoted by Creative Commons, which is also ideal for describing copyright in RDF with the properties cc:permits, cc:requires, cc: prohibits, cc:jurisdiction, cc:legalcode, cc:deprecatedOn; and the classes Work, License, Jurisdiction, Permission, Requirement, and Prohibition[23].

The digital management of rights can be performed by using the Open Digital Rights Language (ODRL)[24].

Education videos can be described by the IEEE Learning Object Metadata (LOM)[25].

Several metadata associated with videos are supported by multiple formats such as personal data on the staff (DC, FOAF, DOAC, Bio, etc.) or geographic positions (DC, Geo[26], etc.).

Due to the overlapping vocabularies of metadata technologies, certain formats have never been widely distributed (e.g., SOMA[27]).

4. COMBINING METADATA

If no additional constraints apply, metadata can be arbitrarily mixed through multiple independently developed vocabularies. For example, microformats can be combined with (X)HTML5 microdata. Another example is Dublin Core that can be used simultaneously with LOM or ODRL metadata[28]. Depending on the metadata formats, duplication might be eliminated by transforming metadata from one format to another (e.g., microformats to the far more generic RDFa[29]).

5. VIDEO METADATA IN SEO

Apart from special approaches of search engines for labeling videos such as the Video Sitemap of Google[30], standard metadata formats can also be provided in a machine-readable way. Two examples are described in the following sections.

5.1 Facebook Share and RDFa–Rich Snippets

The image and video resource URLs are required for Facebook Share (image_src and video_src). The medium property supports the values audio, image, video, news, blog and mult. Video size can be provided by using the video_width and video_height properties. The MIME type of videos can be identified by video_type (with the value application/x-shockwave-flash). A brief description of up to 200 characters can be written using the description property. The title of the video, that can be a maximum of 60 characters, can be added by the title property, i.e.

```
<meta name=”title” content=”Smith plays  BWV543" />
<meta name=”description” content=”Organist John
  Smith plays Praeludium and Fuge in A minor by
  J. S. Bach” />
<link rel=”image_src” href=”http://
  example.com/543thumb.jpg” />
<link rel=”video_src” href=”http://example.com/bach/
  543vid.swf” />
<meta name=”video_width” content=”640" />
<meta name=”video_height” content=”385" />
<meta name=”video_type” content=”application/
  x-shockwave-flash” />
```

All these properties are identified by Google.

5.2 Yahoo! SearchMonkey RDFa

Yahoo! Searchmonkey metadata can be provided on the object tag as follows:

```
<object type=”application/x-shockwave-flash”
  width=”480" height=”385"
  data=”http://www.youtube.com/v/a38-
  oj8VEXI&amp;hl=en_US&amp;fs=1&amp;”
  rel=”media:video” resource=”http://
  www.youtube.com/v/a38-
  oj8VEXI&amp;hl=en_US&amp;fs=1&amp;”
  xmlns:media=”http://search.yahoo.com/
  searchmonkey/media/”
  xmlns:dc=”http://purl.org/dc/terms/”>
<a rel=”media:thumbnail” href=”http://
  example.com/prev.jpg” />
```

The media namespace xmlns:media is required and the only acceptable value is “http://search.yahoo.com/searchmonkey/media/”. The GIF, JPEG, or PNG image with a resolution of 105 x 93 pixels that previews the video before the user clicks on the play button should be defined by the URI as the href attribute value of media:thumbnail. The video to be played when the user clicks the play button should be defined the resource of media:video. All other tags are optional, including the Dublin Core namespace (xmlns:dc) and DC metadata (dc:date, dc:creator, dc:subject, dc:identifier, dc:license, dc:contributor, dc:description), the media metadata (media:title, media:height, media:width, media:duration, media:player, media:type, media:region, media:views), as well as review:rating[31].

6. EMBEDDING VIDEOS AS FLASH

HTML5 has re-introduced the embed element for embedded contents that require plugins. Its name is identical to those of the HTML element used prior to XHTML *object* elements; however, the embed element has been modified in HTML5. The standard-compliant and browser-independent method for Flash embedding is Flash Satay which is also suggested by the World Wide Web Consortium[32]. The reason for the popularity of this method is the pre-installed Flash Player supported by most browsers. However, videos should be embedded as videos rather than generic or Flash objects.

7. LIMITATIONS

The advanced markup element of HTML5 dedicated to videos comes with a lack of consensus about codecs to support. To avoid a barrier for those videos already published in a special format, standardisation would be required that considers not only the popularity of existing codecs, but also the emerging formats for high definition videos.

In spite of the large potential of video metadata, it cannot be fully exploited until social media and video sharing portals either remove all embedded metadata during upload, or apply a new, on-the-fly generated file without such metadata (even in another file format).

8. CONCLUSIONS

The paper provides an investigation of semantic elements as well as generic and specific video metadata standards that can serve as the basis for standardisation.

The newly introduced video element of HTML5 is a promising solution for publishing video contents in a machine-processable form.

Apart from the self-closing tags in XHTML5, metadata embedding is similar in HTML5 and its XML serialization, but there is a huge difference when it comes to namespaces. Due to the overlapping vocabularies of video metadata formats, the most comprehensive ones should be selected and combined. (X)HTML5 therefore supports a variety of metadata formats and provides one of its own, but most metadata can be substituted with the Resource Description Framework in attributes. Due to the enormous number of video contents and streaming media, however, the standardiSation of a more specific metadata format is desired for online videos and movie clips.

REFERENCES

1. Hunter, Jane & Armstrong, Liz. A comparison of schemas for video metadata representation. *Computer Networks* 1999, **31**(11-16), 1431-451.

2. Cuggia, Marc; Mougin, Fleur & Le Beux Pierre. Indexing method of digital audiovisual medical resources with semantic web integration. *Int. J. Med. Inf.,* 2005, **74**(2), 169-77.

3. Gómez-Romero, Juan; Patricio, Miguel A.; García, Jesús & Molina, José M. Ontology-based context representation and reasoning for object tracking and scene interpretation in video. *Exp. Syst. App.,* 2011, **28**(6), 7494-510.

4. Brostow, Gabriel J.; Fauqueur, Julien & Cipolla, Roberto. Semantic object classes in video: A high-definition ground truth database. *Patt. Rec. Lett.,* 2009, **30**(2), 88–97.

5. Gayo, Jose Emilio Labra; Ordóñez de Pablos, Patricia & Lovelle, Juan Manuel Cueva. WESONet: Applying semantic web technologies and collaborative tagging to multimedia web information systems. *Com. Hum. Behav.,* 2010, **26**(2), 205-09.

6. Küçük, Dilek & Yazýcý, Adnan. Exploiting information extraction techniques for automatic semantic video indexing with an application to turkish news videos. *Knoweldge-based Syst.,* 2011 (in Press).

7. World Wide Web consortium. The SMIL 3.0 metainformation module. http://www.w3.org/TR/SMIL/smil-metadata.html#smilMetadataNS-metaModule (accessed on 9 April 2011).

8. World Wide Web Consortium. The video element. *In* HTML5. A vocabulary and associated APIs for HTML and XHTML. http://www.w3.org/TR/html5/video.html#video (accessed 09 April 2011).

9. The internet assigned numbers authority. MIME Media Types. http://www.iana.org/assignments/media-types/. (accessed on 1 January 2011).

10. Internet engineering task force. Multipurpose internet mail extensions (MIME) Part One: Format of Internet message bodies. RFC 2045. http://tools.ietf.org/html/rfc2045 (accessed on 18 February 2011).

11. The Internet society. PostScript subtype. *In* Multipurpose internet mail extensions (MIME) Part Two: Media Types. RFC 2046. http://tools.ietf.org/html/rfc2046 (accessed on 1 January 2011).

12. Internet engineering task force. MIME type registration for MPEG-4. RFC 4337. http://tools.ietf.org/html/rfc4337 (accessed on 02 February 2011).

13. Internet engineering task force. Ogg media types. RFC 5334. http://tools.ietf.org/html/rfc5334 (accessed on 2 January 2011).

14. Internet assigned numbers authority. Registration of the MIME content-type/subtype video/quicktime. http://www.iana.org/assignments/media-types/video/quicktime (accessed on 2 January 2011).

15. Microsoft corporation. MIME type settings for windows media services. http://support.microsoft.com/kb/288102 (accessed on 2 January 2011).

16. World Wide Web Consortium. MIME types. *In* HTML5. A vocabulary and associated APIs for HTML and XHTML, 2011. http://www.w3.org/TR/html5/video.html#dom-navigator-canplaytype. (accessed on 9 April 2011).

17. Lewis, Emily P. Microformats made simple. New Riders, Berkeley, CA, 2009.

18. World Wide Web Consortium. RDFa Core 1.1. Syntax and processing rules for embedding RDF through attributes. http://www.w3.org/TR/rdfa-core/ (accessed on 4 April 2011).

19. World Wide Web consortium. HTML Microdata. http://www.w3.org/TR/microdata/ (accessed on 4 April 2011).

20. World Wide Web Consortium. Introduction. *In* RDF/XML syntax specification. http://www.w3.org/TR/rdf-syntax-grammar/#section-Introduction (accessed on 10 April 2011).

21. World Wide Web Consortium. HTML+RDFa 1.1. Support for RDFa in HTML4 and HTML5. http://www.w3.org/TR/rdfa-in-html/ (accessed on 4 April 2011).

22. Dublin core metadata initiative. dublin core metadata element set, Version 1.1. http://dublincore.org/documents/dces/ (accessed on 8 April 2011).

23. Creative commons. describing copyright in RDF creative commons rights expression language. http://creativecommons.org/ns (accessed on 21 February 2011).

24. World Wide Web Consortium. Open Digital Rights Language (ODRL) Version 1.1. http://www.w3.org/TR/odrl (accessed on 21 February 2011).

25. IMS global learning consortium. IMS meta-data best practice guide for IEEE 1484.12.1-2002 standard for learning object metadata, Version 1.3. http://www.imsglobal.org/metadata/mdv1p3/imsmd_bestv1p3.html (accessed on 21 February 2011).

26. World Wide Web consortium. WGS84 geo positioning: an RDF vocabulary. http://www.w3.org/2003/01/geo/wgs84_pos.rdf. (accessed on 21 February 2011).

27. SOMA Group. SOMA Metadata Element Set. http://soma-dev.sourceforge.net/SOMA_Metadata_1.htm. (accessed on 8 April 2011).

28. Open digital rights language initiative. odrl 1.1 expression language schema. http://odrl.net/1.1/ODRL-EX-11-DOC/index.html (accessed on 23 October 2010).

29. Adida, Ben. hGRDDL: Bridging microformats and RDFa. Web Sem. 2008, **6**(1), 54-60.

30. Google Incorporation. Creating a video sitemap. http://www.google.com/support/webmasters/bin/answer.py?answer=80472 (accessed on 10 April 2011).

31. Yahoo! Incorporation. SearchMonkey–Video. http://developer.search.yahoo.com/help/objects/video (accessed on 15 October 2010).

32. World Wide Web consortium. How can I include Flash in valid (X)HTML Web pages? *In* Help and FAQ for the Markup Validator. http://validator.w3.org/docs/help.html#faq-flash (accessed on 15 December 2010).

About the Author

Dr Leslie F. Sikos is a Web standardista and metadata expert. He is a Computer Scientist and holds a PhD in Information Technology. He is a senior Web developer interested mainly in Web Quality Assurance including, but not limited to, website standardisation, the Semantic Web, Web accessibility, and multimedia. He has written several influential books and is an IT professional also known for IT-related activities such as videography.