INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC1/SC29/WG11
CODING OF MOVING PICTURES AND AUDIO

Source	**Telecom ParisTech, CableLabs**
Status	**For consideration during the 108th MPEG meeting**
Title	**Mapping of DASH Role to HTML 5**
Author	Cyril Concolato (Telecom ParisTech), Jean Le Feuvre (Telecom ParisTech), Bob Lund (CableLabs)

1   Introduction

The HTML 5 specification defines how to expose media tracks in the user interface of media players depending on the type of the tracks (audio, video, text) but also on the notion of "kind" [1][2]. The HTML 5 specification, following recommendations from the 'in-band tracks community group' [3], proposes a mapping of DASH Role elements to the kind attribute. However, this mapping is not complete as some values are not defined DASH. This contribution proposes to define new values for the Role descriptor.

2   HTML 5 kinds and mapping to DASH

The following tables (extracted from the HTML 5 specification) indicate the mapping of DASH Role elements to the HTML kind values for the different types of tracks. The red highlights indicate mappings which are not yet possible.

Text track kinds	Semantics	Corresponding DASH Role
subtitles	Transcription or translation of the dialogue, suitable for when the sound is available but not understood (e.g. because the user does not understand the language of the media resource's audio track). Overlaid on the video.	"subtitle"
captions	Transcription or translation of the dialogue, sound effects, relevant musical cues, and other relevant audio information, suitable for when sound is unavailable or not clearly audible (e.g. because it is muted, drowned-out by ambient noise, or because the user is deaf). Overlaid on the video; labeled as appropriate for the hard-of-hearing.	"caption"
descriptions	Textual descriptions of the video	Not defined

	component of the media resource, intended for audio synthesis when the visual component is obscured, unavailable, or not usable (e.g. because the user is interacting with the application without a screen while driving, or because the user is blind). Synthesized as audio.	
chapters	Chapter titles, intended to be used for navigating the media resource. Displayed as an interactive (potentially nested) list in the user agent's interface.	Not defined
metadata	Tracks intended for use from script. Not displayed by the user agent.	Not defined

From the above table, we suggest defining the following new values for the Role element:
- "description" to identify text tracks to be used by audio synthesis engines.
- "chapter" to identify text tracks to be used to for providing chaptering, typically a WebVTT track.

Regarding the "metadata" kind, in some cases, for instance when metadata is carried in ISOBMFF metadata tracks, simply looking at the codecs parameter of the representation to find the 'metx' or 'mett' value can help an application to determine that the kind is "metadata". In other cases, when no codecs parameter is exposed (for instance when using a WebVTT file, carrying metadata, a WebVTT of type metadata in an MP4 file), it is not so straightforward. In such cases, a "metadata" role could be used. An alternative is to decide that the kind is "metadata" based on the absence of Role elements of type "chapter", "description", "subtitle" or "caption". If this is acceptable to the group, therefore no new role is necessary.

Audio/Video track kinds	Semantics	Corresponding DASH Role
alternative	A possible alternative to the main track, e.g. a different take of a song (audio), or a different angle (video).	"alternate" without "main" and "commentary" roles, and, for audio, without the "dub" role (other roles ignored).
captions	A version of the main video track with captions burnt in. (For legacy content; new content would use text tracks.)	"caption" and "main" roles together (other roles ignored).
descriptions	An audio description of a video track.	Not defined
main	The primary audio or video track.	main" role without "caption", "subtitle", and "dub" roles (other roles ignored).
main-desc	The primary audio track, mixed with audio descriptions.	Not defined
sign	A sign-language interpretation of an audio track.	Not defined

subtitles	A version of the main video track with subtitles burnt in.	"subtitle" and "main" roles together (other roles ignored).
translation	A translated version of the main audio track.	"dub" and "main" roles together (other roles ignored).
commentary	Commentary on the primary audio or video track, e.g. a director's commentary.	"commentary" role without "main" role (other roles ignored).

From the above table, we suggest to define the following new values for the Role element:
- "description" (as proposed above for text tracks).
- "sign" to identify video tracks to be used to provide signed language.

The "main-desc" kind can be determined by the combination of the "main" and "description" Roles. Therefore no new role is necessary.

3   Conclusion

We recommend MPEG to amend the DASH specification to include the proposed new roles: "description", "chapter", "sign".

4   References

[1]   http://www.w3.org/html/wg/drafts/html/master/embedded-content.html#attr-track-kind-subtitles
[2] http://www.w3.org/html/wg/drafts/html/master/embedded-content.html#dom-audiotrack-kind
[3] http://www.w3.org/community/inbandtracks/wiki/Main_Page