**INTERNATIONAL ORGANISATION FOR STANDARDISATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC1/SC29/WG11**

**CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC1/SC29/WG11 MPEG2016/N17837**

**July 2018, Ljubljana, SLO**

|  |  |
| --- | --- |
| **Source** | **Communications** |
| **Status** | **Output document** |
| **Title** | **Draft of White Paper on Media Linking Application Format** |
| **Author** | Alberto MESSINA, Leonardo CHIARIGLIONE |

# MPEG-A Part 18: Media Linking Application Format

ISO/IEC 23000 Part 18, also known as “Media Linking Application Format” (in short MLAF) is a standard aimed at providing a common framework for media content interlinking. As part of the MPEG-A family of standards, MLAF selected readily tested and verified technologies taken from the MPEG corpus of standards and combined them to form a new kind of Multimedia Application Format (MAF).

# Rationale for and Development of the MLAF standard

The development of MLAF has been prompted by existing many examples of services where media transmitted for consumption on a primary device give hints to users to consume related media on a secondary or companion device. To facilitate interoperability of such services, it is beneficial to define a data structure (a “format”) that codifies the relationship between the two information sources. Starting from this sample use case, the development of the standard finally delivered a generic mechanism to link content items and to graphically represent the links.

The development of the MLAF standard has stemmed from the experience of the BRIDGET project [1], an European Commission’s funded initiative under the auspices of Framework 7, active between 2013 and 2016.

# The Key concept: the *bridget*

The core conceptualization of MLAF is a data structure called “bridget”. A bridget is a link between a “source” content and a “destination” content. The bridget contains information on the source content and on the destination content, as well as on the link between the two and on how the information contained in the bridget is to be presented to users consuming the source content in order to enable them to make considerate decisions about whether to consume the destination content.

# Fundamental MLAF technology

MLAF is subdivided in two main chapters: 1) MLAF representation, i.e. technologies aimed at representing the bridgets, together with descriptive information related to the linked items and to the link itself; 2) MLAF presentation, i.e. the technologies aimed at providing a reference rendering framework for bridgets and a way to reference other means to render the same information.

The first part (MLAF representation), is based on MPEG-21 Digital Item Declaration [2]. The second part (MLAF presentation), is based on MPEG-4 Scene description and application engine [3]. MLAF makes the best usage of both referenced standards by restricting their generic capability to cover the underlying business case.

## MLAF representation

The structure of MLAF representation is based on MPEG-21 Digital Item Container. In particular, the top level structure, named the MLAF Container, contains a reference to the Source Item, i.e. the Digital Item which acts as source content from which the links starts, and a set of bridgets representing the links (Figure 1).

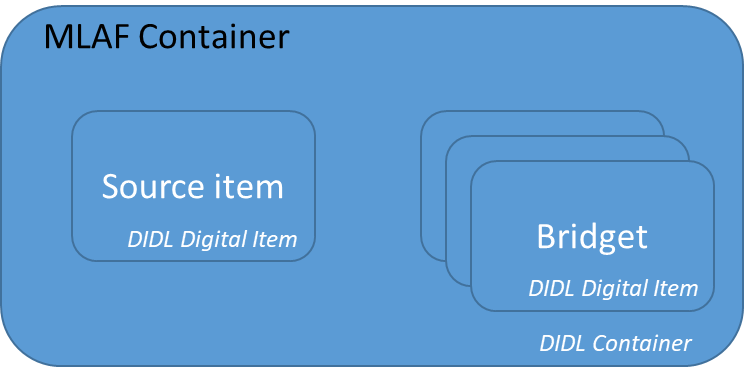


Figure 1. Top-level MLAF Container structure.

A bridget is another specialization of MPEG-21 Digital Item. It specifies a link towards one or more destination content items and information about the spatio-temporal scope of the source item for which the link applies, implemented as a specialized MPEG-21 Annotation (Figure 2). Let’s make an example to illustrate a typical situation. If an author wants to associate a textual description to a temporal segment of a video, then we will have:

* A main MLAF Container containing a description of the video (the Source Item) and one bridget;
* Inside the bridget, the boundaries of the temporal segment of the video to which to associate the textual content (the Structural Annotation);
* The textual content (the Destination Item).

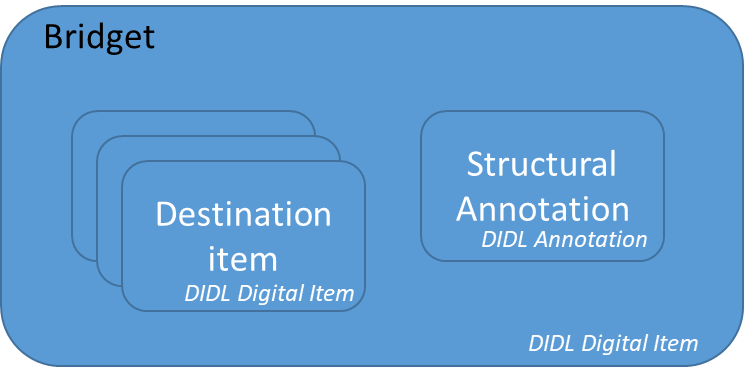


Figure 2. The Bridget structure.

The spatio-temporal scope inside the specialized MPEG-21 Annotation is expressed either through an MPEG-7 MovingRegionSpatioTemporalDecompositionType or through an MPEG-7 AudioVisualSegmentType.

With this high-level powerful syntax, the author, through appropriate user tools, is able to specify a wide range of possible associations and be as precise and granular as possible exploiting the expressive power of these two MPEG-7 tools and the general descriptive capability of MPEG-21 Digital Item.

## MLAF presentation

TBA

# Benefits provided by MLAF

Thanks to MLAF, what is foreseeable is a sort of “layered” approach at producing bridgets in which actors with different roles define bridgets under different perspectives and possibly concurring at the same time (Figure 3). Authors of audiovisual content (e.g., television programmes) will define bridget end points (i.e. sources and destination content items) following the criteria matching with the editorial intention, main distribution channel or target audience of the programme. At the same time, marketing and commercial operators (e.g. advertisement agents) will define such end points following their own mind-setting, which may be independent from the authorial perspective. Last, but definitely not least, final users can define their own ways for bridgets through social media interaction. All the above approaches can include not only the generation of the linking information but also of information related to how referenced content have to be presented graphically or should interact with the user.

23000-18_ed1fig1_e

Figure 3. Bridget creation ecosystem.

# Usage of MLAF

## MLAF as an authoring format

Thanks to its rich and comprehensive descriptive toolset, MLAF can be fruitfully used as an authoring format by users interested in authoring link information among content items of a collection or between their content items and external data sources. The inherent extensibility of MPEG-21 Digital Item Declaration technology makes it easy to devise authoring tools fulfilling specific domain requirements (e.g., identification mechanisms, classification taxonomies).

## MLAF as a presentation format

MLAF carries a native format for presenting bridget information, which is based on MPEG-4 Scene description and application engine. Nevertheless, a bridget can be directly linked to any external presentation/rendering resource (e.g., an HTML page, an SVG graphics or other),

## MLAF as an exchange format

MLAF is essentially an XML dialect. This makes it natively apt to be used as an information exchange format between authors and users of bridgets.

# References

1. BRIDGET, <http://ict-bridget.eu>
2. ISO/IEC 21000-2:2005, “Digital Item Declaration”.
3. ISO/IEC 14496-11:2015, “Scene description and application engine”