INTERNATIONAL STANDARD

ISO/IEC 23000-18

First edition
2018-05

Information technology — Multimedia application formats (MPEG-A) —

Part 18: Media linking application format

Technologies de l'information — Formats des applications multimédia (MPEG-A) —
Partie 18: Format des applications de liaison de médias

Reference number
ISO/IEC 23000-18:2018(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of document should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

A list of all parts in the ISO/IEC 23000 series can be found on the ISO website.

Introduction

ISO/IEC 23000 (also known as “MPEG-A”) is an MPEG standard that supports a fast track to standardization by selecting readily tested and verified technologies taken from the MPEG corpus of standards and combining them to form a Multimedia Application Format (MAF). If a needed piece of technology is not available from the said corpus, then additional technologies originating from other organizations may also included by reference in order to facilitate the envisioned MAF.

The development of the MAF called “Media Linking Application Format” (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.

The core of MLAF structure is the representation of the structure of bridgets. A bridget is a pivotal information element that links Source Content and Destination Content. A bridget is associated to its Presentation Information. Clause 7 specifies the representation of bridgets. Clause 8 specifies the presentation of bridgets.

The XML Schema of MLAF is defined in Annex A. Technology for bridget Presentation based on XMT is included in Annex B. Bridget presentation in BIFS is defined in Annex C. Annex D reports some examples.

Information technology — Multimedia application formats (MPEG-A) —

Part 18:
Media linking application format

1 Scope

This document specifies 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, 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.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 14496-11:2015, *Information technology — Coding of audio-visual objects — Part 11: Scene description and application engine*

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— IEC Electropedia: available at http://www.electropedia.org/

— ISO Online browsing platform: available at http://www.iso.org/obp

3.1 MPEG-4 terminology

3.1.1
audiovisual scene
AV scene
set of audiovisual objects together with scene description information that defines their spatial and temporal attributes including behaviours resulting from object and user interactions

3.1.2
binary format for scene
BIFS
coded representation of a parametric scene description format

3.2 MPEG-7 terminology

3.2.1
segment
spatial or temporal unit of multimedia, for example, a temporal segment of video, or a segment of an image

3.2.2
audiovisual segment
element of MPEG-7 multimedia description schemes structure description tools describing a time-continuous arbitrary part (segment) of an audiovisual media

Note 1 to entry: See ISO/IEC 15938-5.

3.2.3
region
spatial unit of multimedia, for example, a 2D spatial region of an image, or a *moving region* (3.2.4) of video

3.2.4
moving region
region (3.2.3) changing its shape during a video timeline

3.3 MPEG-21 terminology

3.3.1
digital item identification
DII
standard used by MPEG-21 for identification of digital item and their *components* (3.3.5)

Note 1 to entry: See ISO/IEC 21000-3.

3.3.2
resource
individually identifiable asset such as a video or audio clip, an image or a text

Note 1 to entry: All resources shall be locatable via an unambiguous address.

3.3.3
container
structure that allows *items* (3.3.4) and/or containers to be grouped

Note 1 to entry: A container itself is not an item; containers are groupings of items and/or containers. Descriptors allow for the "labelling" of containers with information appropriate for the purpose of the grouping.

3.3.4
item
grouping of sub-items and/or *components* (3.3.5) that are bound to relevant *descriptors* (3.3.6)

Note 1 to entry: These descriptors contain information about the item. Items may contain choices, which allow them to be customized or configured. Items may be conditional (on predicates asserted by selections defined in the choices). An item that contains no sub-items can be considered a whole. An item that does contain sub-items can be considered a compilation. Items may also contain annotations to their sub-parts.

3.3.5
component
binding of a *resource* (3.3.2) to a set of *descriptors* (3.3.6)

Note 1 to entry: These descriptors are information concerning all or part of the specific resource instance. A component itself is not an item; components are building blocks of items. Components may be conditional (see ISO/IEC 21000-3 for details).

3.3.6
descriptor
associates information with the enclosing entity

Note 1 to entry: This information may be a component (such as a thumbnail of an image, or a text component) or a textual statement. Descriptors may be conditional (see ISO/IEC 21000-3 for details).

3.3.7
statement
literal textual value that contains information, but not an asset

EXAMPLE Examples of likely statements include descriptive, control, revision tracking or identifying information (such as an identifier as described in ISO/IEC 21000-3).

3.3.8
annotation
describes a set of information about another identified entity of the model without altering or adding to that entity

Note 1 to entry: The information can take the form of assertions, descriptors and anchors.

3.3.9
anchor
binds *descriptors* (3.3.6) to a *fragment,* which corresponds to a specific location or part of a *resource* (3.3.2)

Note 1 to entry: These descriptors are information concerning all or part of the fragment. Anchors may be conditional.

3.3.10
condition
describes the enclosing entity as being optional, and links it to the *selection*(s) that affect its inclusion

Note 1 to entry: Multiple predicates within a condition are combined as a conjunction (an AND relationship). Any predicate may be negated within a condition. Multiple conditions associated with a given entity are combined as a disjunction (an OR relationship) when determining whether to include the entity.

3.4 Terminology inherited from other standards — EBU Core terminology

3.4.1
core metadata
core descriptive information regarding a *resource* (3.3.2)

Note 1 to entry: Core descriptive information include, e.g., titles, contributors, genres, technical formats of a given resource.

3.4.2
metadata provider
entity (person or organization) providing metadata

3.5 MLAF terminology

3.5.1
bridget
link from a piece of *source content* (3.5.2) (e.g. a TV programme) to a piece of *destination content* (3.5.3) (e.g. an interactive media element such as a photograph, a multimedia presentation, a *segment* (3.2.1) or another TV programme, an application)

3.5.2
source content
content that is linked to *destination content* (3.5.3) through a *bridget* (3.5.1)

Note 1 to entry: The source content is the content from which the bridget originates.

3.5.3
destination content
content that is linked to the *source content* (3.5.2) through a *bridget* (3.5.1)

Note 1 to entry: The destination content is the content to which the bridget points.

3.5.4
spatiotemporal scope
information that describe the spatiotemporal *region* (3.2.3) of the linked *source content* (3.5.2) or of the linked *destination content* (3.5.3) that are involved in the link provided by a *bridget* (3.5.1)

3.5.5
global bridget
bridget (3.5.1) which is active for the whole duration of the *source content* (3.5.2)

3.5.6
spatiotemporal-aligned bridget
bridget (3.5.1) which is active for a determined *spatiotemporal scope* (3.5.4) of the *source content* (3.5.2)

4 Abbreviated terms

BIFS	BInary Format for Scenes
DID	MPEG-21, Digital Item Declaration
DIDL	MPEG-21, Digital Item Declaration Language
DII	MPEG-21, Digital Item Identification
MLAF	Media Linking Application Format
XMT	Extensible MPEG-4 Textual Format

5 MLAF model

5.1 General

Bridgets are links which exist because of some inherent semantic relationship between content items. As such, they can be products of an editorial decision, taken by someone as the result of the inspection (which can be manual or automatic) of content items, and can be products of a workflow which involves different roles taking care of finding, organizing and finally crafting the data that constitute them. The nature of a bridget is, however, quite different than traditional linearity of media content, and as such it induces a different, more “distributed” workflow. In fact, whether a piece of media content is a candidate source or destination for a bridget can be the result of an editorial decision taken at any moment and by quite different kind of users.

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. Authors of 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.

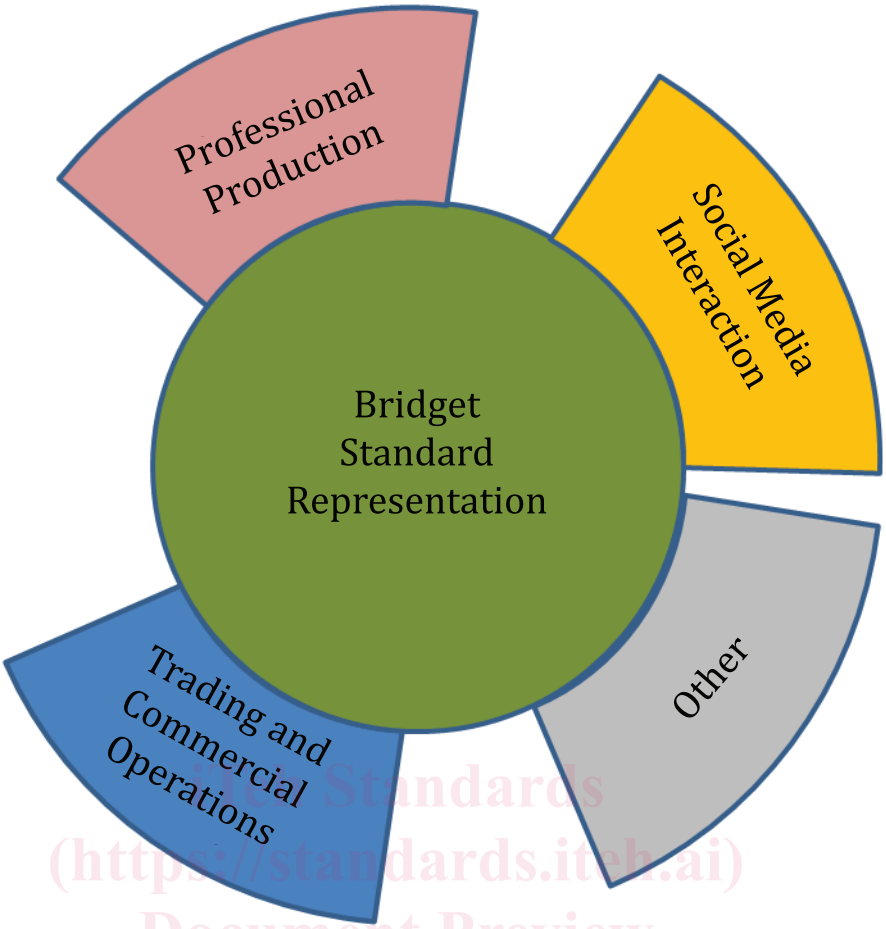

Figure 1 — Bridget creation workflow

The result is that the way in which bridget information would be created is quite different from traditional linear approaches at media production, and more following a daisy chain paradigm (see Figure 1).

Therefore, a standard format for representing and exchanging bridget-related information fosters integration of all those systems having a role in generating bridget information in the different and heterogeneous aforementioned domains. This document specifies this standard format as Media Linking Application Format (MLAF).

5.2 Functional requirements

a) Source and Destination Content of a bridget can be

   1) a file or a stream, or

   2) a static medium (e.g. an image) or a dynamic medium (e.g. a video).

b) A bridget shall include a unique and persistent identifier.

c) A bridget may include authoring information (e.g. title, date, copyright, etc.).

d) A bridget shall include identification of its Source Content.

e) A bridget shall include information about the link between Source and Destination Content.

   1) A bridget shall include reference to

      i) a specific time instant of a dynamic medium representing the source or destination content, or

      ii) a specific time period of a dynamic medium representing the source or destination content, or

      iii) the entire duration of a dynamic medium representing the source or destination content.

   2) A bridget may include reference to

      i) a specific spatial region of a static medium representing a destination content, or

      ii) a specific spatiotemporal region (e.g. a 2D or 3D object) of a dynamic medium representing a destination content.

   3) A bridget may include metadata describing the context of the link between source and destination content (e.g. the semantics of the relation between source and destination).

f) A bridget may include metadata describing the Source Content:

   1) title, date, copyright, etc.;

   2) content type, media duration, MIME type, file size, etc.

g) A bridget may include textual descriptions of the Source Content.

h) A bridget shall include identification of Destination Content.

i) A bridget may include metadata describing the Destination Content:

   1) title, date, copyright, etc.;

   2) content type, media duration, MIME type, file size, etc.

j) A bridget may include textual descriptions of the Destination Content.

k) A bridget may include information on how a bridget should be presented.

l) A bridget may include an icon or image to visually present the bridget to the user.

m) A bridget may include information on visual/acoustic presentation of the Destination Content.

n) A bridget may include consumption-related information, e.g.

   1) parental rating;

   2) accessibility;

   3) device requirements (e.g. stereo headphones);

   4) target users.

o) A bridget may include rights and payment information for destination content consumption.

p) A bridget may include information about the service providing the destination content.

