International Standard

ISO 16684-4

Graphic technology — Extensible metadata platform (XMP) specification —

Part 4: Use of XMP for semantic units

First edition
2024-01

Reference number
ISO 16684-4:2024(en)

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

Published in Switzerland

Contents

Page

Foreword ..... **iv**
Introduction ..... **v**
1 Scope **1**
2 Normative references **1**
3 Terms and definitions **1**
4 Semantic unit **3**
4.1 General 3
4.2 Types of semantic units 3
5 Defining and referencing semantic unit **5**
5.1 Defining semantic units with target resources 5
5.2 Required Keywords 5
5.3 Simple valued XMP: selecting a target source resource 6
5.4 Array valued XMP: selecting a set of specific resources 6
5.4.1 General 6
5.4.2 Array values 6
5.4.3 Structure valued XMP properties 6
5.5 Identifying and dereferencing semantic unit 6
6 Serialization **6**
Annex A (informative) **Use case – Digitized books** **8**
Annex B (informative) **Difference between XMP and the Web Annotation Model** **12**
Bibliography **13**

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

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 ISO 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).

ISO draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO takes no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO had not received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents. ISO shall not be held responsible for identifying any or all such patent rights.

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

For an explanation of 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 www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 171, *Document management applications*, Subcommittee SC 2, *Document file formats, EDMS Systems and authenticity of information.*

A list of all parts in the ISO 16684 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

Traditional concepts and uses of metadata have been applied to describe a file or a collection, but this approach cannot meet the needs of flexible, sub-file and across-file level information exchange and sharing. Researchers, publishers, readers and machines require new approaches to describe content in flexible and multi-faceted ways for data sharing and data exchange (e.g. chapter, figure, image, table, formula). Specifically, in the linked data (either open or closed) web environment, any piece of information can be described, referenced and linked with or without relationships with any other data, objects, and/or files.

By implementing this document, textual and non-textual content can be described, used and shared at an atomic level, file level and across-file level. This information can be used for multiple purposes, including content access, digital preservation, scientific research and publishing. Recent developments in computer vision and machine learning make it possible to create, link and capture semantics from documents, videos and audios. Machines can learn and create meaningful metadata using a variety of artificial intelligence (AI) machine learning models. This enhances existing resources description and creates new opportunities in content generation, content use and content re-use. For instance, the general public benefits from semantic-rich content through enhanced knowledge access, discovery and integration (e.g. see the use cases in the Annex A), while scholars can utilize semantic-rich content for content discovery, integration and scholarly communication. For business, the availability of semantic-rich content creates new opportunities for lower cost, higher productivity, and better user satisfaction. Machines can also utilize semantic-rich content and metadata for a variety of purposes.

Graphic technology — Extensible metadata platform (XMP) specification —

Part 4:
Use of XMP for semantic units

1 Scope

This document:

a) introduces the concept of the semantic unit (SU);

b) provides requirements and guidance on how to define the target resource(s) in an SU by adopting the "target" syntax from the Web Annotation Model;

c) provides requirements and guidance on the extensible metadata platform (XMP) serialization syntaxes for SU.

This document broadens the concept of XMP specified in ISO 16684-1 so that XMP can be used to describe an SU. A new flexible way of defining and describing SUs aims to bring innovation to textual and non-textual content, metadata, linked data, big data and artificial intelligence.

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 16684-1:2019, *Graphic technology — Extensible metadata platform (XMP) — Part 1: Data model, serialization and core properties*

ISO 16684-3:2021, *Graphic technology — Extensible metadata platform (XMP) specification — Part 3: JSON-LD serialization of XMP*

Web Annotation Data Model *W3C Recommendation,* 2017, https://www.w3.org/TR/annotation-model/

3 Terms and definitions

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

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

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

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

3.1
audio
frequency corresponding to a sinusoidal sound wave audible to the normal human ear (from about 16 Hz to 16 kHz)

[SOURCE: IEC 60050-702: 1992, 702-01-08]

Note 1 to entry: Audio content is primarily intended to be heard. Alternatively, it can be described by other expressions such as text, image, and video.

3.2
image
stored description of a graphic picture along with metadata, which is primarily intended to be seen

Note 1 to entry: Alternatively, it can be described by other content expressions such as text, audio, and video.

3.3
dataset
stored description of sets of data along with metadata, which is primarily intended to be processed by software

3.4
external resource
resource which is available outside of the current one

Note 1 to entry: It is de-referenceable from its Internationalized Resource Identifier (IRI).

3.5
selector
method used to select and describe the desired segment(s) from the source resource(s)

Note 1 to entry: The nature of a selector will be dependent on the types of resource, as the methods to select segments from various media-types will differ. Multiple selectors can be given to describe the same segment in different ways.

Note 2 to entry: See W3C Recommendation: 2017, 4.2 for further information.

3.6
semantic unit
SU

user-defined piece of information (content) consisting of one or more target(s) resources, which is logically, semantically or structurally separated from other information and can be meaningfully described as a unit

Note 1 to entry: An SU may be described by one or more extensible metadata platform (XMP) packets.

Note 2 to entry: An SU can be any piece of information (e.g. character, word, phrase, sentence, paragraph, chapter) across any information container (e.g. a file).

Note 3 to entry: An SU can be one of the following representations depending on its content.

— For content interpreted by visual systems (e.g. documents, maps), it can be an area containing content such as characters, figures, tables, images, and formulas.

— For content interpreted by hearing and/or visual systems (e.g. audio, video), it can be the length of starting and stopping time, or data containing such content.

— For content interpreted by smell, it can be the smell data (e.g. location, timestamp, chemical data), and/or interpretations of the smell.

— For content not sensible beyond the five human senses, it can be the data and/or its derivatives.

3.7
target
resource, which is a “specificResource” identified by “source”

Note 1 to entry: See W3C Recommendation: 2017, Section 4 for further information.

3.8
text
sequence of characters or glyphs, or a visual representation thereof

Note 1 to entry: Textual content is primarily intended to be read. Alternatively, it can be expressed by other means such as *audio* (3.1), *image* (3.2), or *video* (3.10) with metadata.

3.9
user
person or machine who defines, describes, and/or uses a *semantic unit* (3.6)

3.10
video
recording of an *image* (3.2) or set of images; content consists of moving images with or without *audio* (3.1)

4 Semantic unit

4.1 General

A semantic unit (SU) is a user-defined content that can be associated with one or more XMP packets. The user can apply this document to encapsulate sufficient information (including semantics) for the SU. ISO 16684-1 specified a way to declare an XMP packet but applies to the entire resource stored as a file.

This document specifies a way to create and to define any resource that can be a piece of information in a file container or cross-file boundary.

4.2 Types of semantic units

An SU can be any piece of information such as zero or more characters, an image, a video frame, a resource on the internet, or a combination of these.

An SU can be a content area inside an image, table, page, file, bitstream, and Internationalized Resource Identifier (IRI). It can enable content cross images, tables, pages, files, bitstreams, and IRIs boundary (e.g. inside a file, outside a file, or a combination of multiple files), and a combination of the above.

An SU can include an ordered or unordered set of resources.

A metadata packet (i.e. XMP packet) for an SU can contain any type of metadata (descriptive, structural, technical, administrative) and data.

An SU can be independent or have relations with other SUs.

Examples include:

— Digitized textual documents (see Annex A);

— Images: SUs defining parts of an image can be separated, connected, overlapped, or covered. Figure 1 shows that multiple SUs can be defined within a single image.

SOURCE: Reference [1], reproduced with permission from the author.

Figure 1 — Multiple SUs in an image

— Several multimedia resources and related resources can be combined as part of one SU. For example, a sheet music fragment (see Figure 2) can be augmented with a video clip of a performance, or a webpage describing the history of the piece.