INTERNATIONAL STANDARD

ISO/IEC
15938-5

First edition
2003-05-15

Information technology — Multimedia content description interface —

Part 5:
Multimedia description schemes

Technologies de l'information — Interface de description du contenu multimédia —

Partie 5: Schémas de description multimédia

ISO IEC

iTeh Standards
(https://standards.iteh.ai)
Document Preview

ISO/IEC 15938-5:2003
https://standards.iteh.ai/catalog/standards/iso/cb17b054-bb45-4baf-8eb7-9ff8d72be241/iso-iec-15938-5-2003

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.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 3.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

ISO/IEC 15938-5 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

ISO/IEC 15938 consists of the following parts, under the general title *Information technology — Multimedia content description interface*:

— *Part 1: Systems*

— *Part 2: Description definition language*

— *Part 3: Visual*

— *Part 4: Audio*

— *Part 5: Multimedia description schemes*

— *Part 6: Reference software*

— *Part 7: Conformance testing*

— *Part 8: Extraction and use of MPEG-7 descriptions*

Annex B forms a normative part of this part of ISO/IEC 15938. Annexes A, C and D are for information only.

Introduction

This standard, also known as "Multimedia Content Description Interface," provides a standardized set of technologies for describing multimedia content. The standard addresses a broad spectrum of multimedia applications and requirements by providing a metadata system for describing the features of multimedia content.

The following are specified in this standard:

- **Description Schemes (DS)** describe entities or relationships pertaining to multimedia content. Description Schemes specify the structure and semantics of their components, which may be Description Schemes, Descriptors, or datatypes.

- **Descriptors (D)** describe features, attributes, or groups of attributes of multimedia content.

- **Datatypes** are the basic reusable datatypes employed by Description Schemes and Descriptors

- **Systems tools** support delivery of descriptions, multiplexing of descriptions with multimedia content, synchronization, file format, and so forth.

This standard is subdivided into eight parts:

Part 1 – Systems: specifies the tools for preparing descriptions for efficient transport and storage, compressing descriptions, and allowing synchronization between content and descriptions.

Part 2 – Description definition language: specifies the language for defining the standard set of description tools (DSs, Ds, and datatypes) and for defining new description tools.

Part 3 – Visual: specifies the description tools pertaining to visual content.

Part 4 – Audio: specifies the description tools pertaining to audio content.

Part 5 – Multimedia description schemes: specifies the generic description tools pertaining to multimedia including audio and visual content.

Part 6 – Reference software: provides a software implementation of the standard.

Part 7 – Conformance testing: specifies the guidelines and procedures for testing conformance of implementations of the standard.

Part 8 – Extraction and use of MPEG-7 descriptions: provides guidelines and examples of the extraction and use of descriptions.

This document specifies the Multimedia Description Schemes (MDS) part of the ISO/IEC 15938 standard. This document specifies the MDS description tools and provides informative examples of descriptions. The normative syntax of the MDS description tools is specified in this document using the Description Definition Language (DDL) and the normative semantics is specified using text.

Information technology — Multimedia content description interface —

Part 5:
Multimedia description schemes

1 Scope

1.1 Organization of the document

This International Standard specifies a metadata system for describing multimedia content. This document specifies the Multimedia Description Schemes (MDS) description tools (Description Schemes, Descriptors, and datatypes) that comprise ISO/IEC 15938-5. The following set of subclauses are provided for each MDS description tool, where optional subclauses are indicated as (optional):

- Syntax: specifies the normative syntax of the description tool using Description Definition Language (DDL) (see ISO/IEC 15938-2).

- Binary syntax (optional): specifies the normative binary representation for the description tool, where relevant. If a binary syntax is not specified, then the binary coding method of ISO/IEC 15938-1 shall apply for the description tool.

- Semantic: specifies the normative semantics of the description tool and each of its components (attributes and elements).

- Informative examples (optional): provides informative examples that illustrate the use or instantiation of the description tool in creating descriptions.

1.2 Overview of Multimedia Description Schemes

The MDS description tools are organized on the basis of functionality as shown in Figure 1.
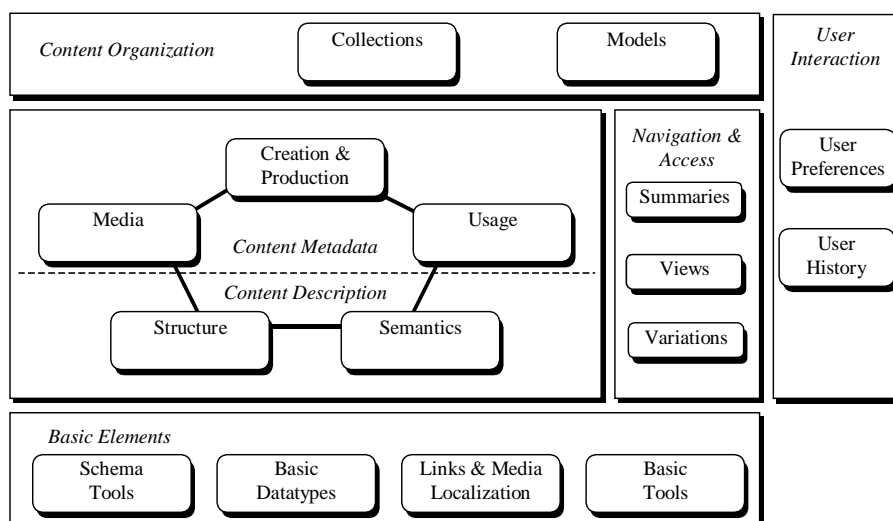

Figure 1 — Overview of the Multimedia Description Scheme (MDS) description tools

The basic elements, which are shown at the bottom level in Figure 1, form the building blocks for the higher-level description tools. The following basic elements are defined: schema tools (the root element, top-level types, description metadata, and packages), basic datatypes (integers, reals, vectors, matrices), linking and media localization tools (spatial and temporal localization), basic description tools (language, text, classification schemes). The basic elements are specified in the following clauses: clause 4 (Schema tools), clause 5 (Basic datatypes), clause 6 (Linking, identification and localization tools), and clause 7 (Basic description tools). Table 1 summarizes the basic elements.

Table 1 — Overview of basic elements

Set of description tools	Functionality
Schema tools (clause 4)	Specifies schema tools that facilitate the making of descriptions. The schema tools include the base type hierarchy of the description tools defined in ISO/IEC 15938, the root element and top-level tools, the multimedia content entity tools, and the package and description metadata tools.
Basic datatypes (clause 5)	Specifies the basic datatypes used by description tools defined in the ISO/IEC 15938.
Linking, identification and localization tools (clause 6)	Specifies the basic datatypes that are used for referencing within descriptions and linking of descriptions to multimedia content.
Basic description tools (clause 7)	Specifies basic tools that are used as components for building other description tools.

The content description tools, which are shown at the middle level in Figure 1, describe the features of the multimedia content and the immutable metadata related to the multimedia content. The following description tools for content description are defined: structure description tools (spatio-temporal segments of multimedia content) and semantic description tools (objects, events). The following description tools for content metadata are defined: media description (storage format, encoding), creation & production (title, creator, classification), and usage (access rights, publication). The content description and metadata tools are related in the sense that the content description tools use the content metadata tools. For example, a description of creation and production or media information can be attached to an individual video or video segment in order to describe the structure and creation and production of the multimedia content. Table 2 summarizes tools for describing content and related metadata.

Table 2 — Overview of tools for describing content and related metadata

Set of description tools	Functionality
Media description tools (clause 8)	Describes the storage of the multimedia data. The media features include the format, encoding, storage media. The tools allow multiple media description instances for the same multimedia content.
Creation and production description tools (clause 9)	Describes the creation and production of the multimedia content. The creation and production features include title, creator, classification, purpose of the creation, and so forth. The creation and production information is typically not extracted from the content but corresponds to metadata related to the content.
Usage description tools (clause 10)	Describes the usage of the multimedia content. The usage features include access rights, publication, and financial information. The usage information may change during the lifetime of the multimedia content.
Structure description tools (clause 11)	Describes the structure of the multimedia content. The structural features include spatial, temporal or spatio-temporal segments of the multimedia content.
Semantics description tools (clause 12)	Describes the "real-world" semantics related to or captured by the multimedia content. The semantic features include objects, events, concepts, and so forth.

The tools for navigation and access, which are shown at the middle level in Figure 1, describe the browsing, summarization, and access of content. These tools are specified in clause 13 (Navigation and access tools). The tools for content organization tools, shown at the top level in Figure 1, describe collections and models of multimedia content. These tools are specified in clause 14 (Content organization tools). The tools for user interaction, shown on the right in Figure 1, describe user preferences pertaining to consumption of multimedia and usage history. These tools are specified in clause 15 (User interaction tools). Table 3 summarizes tools for describing content organization, navigation and access, and user interaction.

Table 3 — Overview of tools for describing content organization, navigation and access, and user interaction

Set of description tools	Functionality
Navigation and access tools (clause 13)	Specifies tools that describe the navigation and access of multimedia. The navigation and access tools describe multimedia summaries and abstracts; describe partitions, views and decompositions of image, video, and audio signals in space, time and frequency; and describe the relationships between different variations of multimedia content.
Content organization tools (clause 14)	Specifies tools that describe the organization and modeling of multimedia content. The content organization tools include collections, probability models, analytic models, cluster models, and classification models.
User interaction tools (clause 15)	Specifies tools related to user interaction with multimedia content. The user interaction tools describe user preferences pertaining to multimedia content, and describe the usage history of users of multimedia content.

Finally, classification schemes, which organize terms that are used by the description tools, are specified in Annex B (Classification Schemes).

2 Normative references

The following normative documents contain provisions which, through reference in this text, constitute provisions of this part of ISO/IEC 15938. For dated references, subsequent amendments to, or revisions of, any of these publications do not apply. However, parties to agreements based on this part of ISO/IEC 15938 are encouraged to investigate the possibility of applying the most recent editions of the normative documents indicated below. For undated references, the latest edition of the normative document referred to applies. Members of ISO and IEC maintain registers of currently valid International Standards. The Telecommunication Standardization Bureau maintains a list of currently valid ITU-T Recommendations.

- ISO 8601, *Data elements and interchange formats — Information interchange — Representation of dates and times*

- ISO 639 (all parts), *Codes for the representation of names of languages*

- ISO 3166-1, *Codes for the representation of names of countries and their subdivisions — Part 1: Country codes*

- ISO 3166-2, *Codes for the representation of names of countries and their subdivisions — Part 2: Country subdivision code*

Note: The current list of valid ISO 3166-1 country and ISO 3166-2 region codes is maintained by the ISO 3166 maintenance agency at the ISO Central Secretariat. Information on the current list of valid region and country codes can be found at http://www.iso.org/mara/iso3166.

- ISO 4217, *Codes for the representation of currencies and funds*

Note: The current list of valid ISO 4217 currency codes is maintained by the official maintenance authority British Standards Institution (http://www.bsi-global.com/iso4217currency).

- XML, *Extensible Markup Language (XML) 1.0,* 6 October 2000 ⟨http://www.w3.org/TR/2000/REC-xml-20001006⟩

- *XML Schema, W3C Recommendation*, 2 May 2001 ⟨http://www.w3.org/XML/Schema⟩

- *XML Schema Part 0: Primer*, W3C Recommendation, 2 May 2001 ⟨http://www.w3.org/TR/xmlschema-0/⟩

- *XML Schema Part 1: Structures*, W3C Recommendation, 2 May 2001 ⟨http://www.w3.org/TR/xmlschema-1/⟩

- *XML Schema Part 2: Datatypes*, W3C Recommendation, 2 May 2001 ⟨http://www.w3.org/TR/xmlschema-2/⟩

- XPath, *XML Path Language, W3C Recommendation*, 16 November 1999 ⟨http://www.w3.org/TR/1999/REC-xpath-19991116⟩

Note: These documents are maintained by the W3C (http://www.w3.org).

- RFC 2045 *Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies*

- RFC 2046 *Multipurpose Internet Mail Extensions (MIME) Part Two: Media Types*

- RFC 2048, *Multipurpose Internet Mail Extensions (MIME) Part Four: Registration Procedures*

- RFC 2045-CHARSETS, *Registered Character set codes of RFC2045*

- RFC 2046-MIMETYPES, *Registered Mimetypes of RFC2046*

Note: The relevant lists can be obtained as follows:

  o MIMETYPES. The current list of registered mimetypes, as defined in RFC2046, RFC2048, is maintained by IANA (Internet Assigned Numbers Authority). It is available from ftp://ftp.isi.edu/in-notes/iana/assignments/media-types/media-types/.

  o CHARSETS. The current list of registered character set codes, as defined in RFC2045 and RFC2048 is maintained by IANA (Internet Assigned Numbers Authority). It is available from ftp://ftp.isi.edu/in-notes/iana/assignments/character-sets.

3 Terms and definitions

3.1 Conventions

3.1.1 Description tools

This part of ISO/IEC 15938 specifies the multimedia description tools as follows:

- **Description Scheme (DS)** – a description tool that describes entities or relationships pertaining to multimedia content. DSs specify the structure and semantics of their components, which may be Description Schemes, Descriptors, or datatypes.

- **Descriptor (D**) – a description tool that describes a feature, attribute, or group of attributes of multimedia content.

- **Datatype** – a basic reusable datatype employed by Description Schemes and Descriptors.

- **Description Tool (or tool)** – refers to a Description Scheme, Descriptor, or Datatype.

3.1.2 Naming convention

In order to specify the multimedia description tools, this part of ISO/IEC 15938 uses constructs provided by the Description Definition Language (DDL) specified in ISO/IEC 15938-2, such as "element," "attribute," "simpleType" and "complexType." The names associated to these constructs are created on the basis of the following conventions:

- If the name is composed of multiple words, the first letter of each word is capitalized, with the exception that the capitalization of the first word depends on the type of construct as follows:

- Element naming: the first letter of the first word is capitalized (e.g. `TimePoint` element of `TimeType`).

- Attribute naming: the first letter of the first word is not capitalized (e.g. `timeUnit` attribute of `IncrDurationType`).

- complexType naming: the first letter of the first word is capitalized, and the suffix "Type" is used at the end of the name (e.g. `PersonType`).

- simpleType naming: the first letter of the first word is not capitalized, the suffix "Type" may be used at the end of the name (e.g. `timePointType`).

Note that when referencing a complexType or simpleType in the definition of a description tool, the "Type" suffix is not used. For instance, the text refers to the "`Time` datatype" (instead of "`TimeType` datatype"), to the "`MediaLocator` D" (instead of "`MediaLocatorType` D") and to the "`Person` DS" (instead of `PersonType` DS).

3.1.3 Documentation convention

The syntax of each description tool is specified using the constructs provided by the DDL specified in ISO/IEC 15938-2, and is presented in this document using a specific font and background as shown in the following example:

```
<complexType name="ExampleType">
    <sequence>
        <element name="Element1" type="string"/>
    </sequence>
    <attribute name="attribute1" type="string" default="attrvalue1"/>
</complexType>
```

The semantics of each description tool is specified in text using a table format, where each row contains the name and a definition of a type, element or attribute as shown in the following example:

Name	Definition
ExampleType	Specifies an ...
element1	Describes the …
attribute1	Describes the …

Diagrammatic notation is sometimes used to depict overviews of the description tools. Such overview diagrams generally depict aggregation and generalization relationships between description tools or the entities being described. In these diagrams, rectangular shapes containing a name denote description tools (DSs or Ds) that are used to describe entities. In some diagrams, multiple description tools are listed in a single rectangle. Large diamond shapes containing a name denote description tools (DSs or Ds) that are used to describe relationships. Furthermore, paths between rectangles or diamonds denote association, generalization or aggregation relationships. Generalization relationships are indicated by a solid path with a filled triangle pointing at the more

general entity. Aggregation (composition) relationships are indicated by a solid path with a (filled) diamond at the aggregating (composing) entity. Aggregation relationship paths may be accompanied by an indication of the multiplicity (minOccurs, maxOccurs) of the relation in text form.

The informative examples are included in separate subclauses, and are presented in this document using a separate font and background as shown in the following example:

```
<Example attribute1="example attribute value">
    <Element1>example element content</Element1>
</Example>
```

Moreover, the schema or set of description tools defined in this document follows a type-centric approach. The description tools are specified by defining the complexType or simpleType for each tool, but not declaring an element of this type at the global scope with the exception of the root element. However, in order to illustrate informative example descriptions in this document, the examples assume that an element of the example type (complexType or simpleType) has been declared, such as a member of another complexType or simpleType. For example, the description above assumes that the following declaration has been made:

```
<element name="Example" type="mpeg7:ExampleType">
```

The term "reserved" is used in specifying the semantics of some description tools. The term "reserved" indicates that particular values are reserved for use in future extensions of ISO/IEC 15938.

3.2 Wrapper of the schema

The description examples and syntax of description tools specified in this document assume that a schema wrapper is provided which identifies the XML Schema namespace (XML Schema) and MPEG-7 namespace:

```
<schema xmlns="http://www.w3.org/2001/XMLSchema"
    xmlns:mpeg7="urn:mpeg:mpeg7:schema:2001"
    targetNamespace="urn:mpeg:mpeg7:schema:2001"
    elementFormDefault="qualified" attributeFormDefault="unqualified">
<!-- Multimedia Description Schemes ISO/IEC 15938-5 -->
```

The following tag is used to close the schema:

```
</schema>
```

3.3 Terminology

For the purposes of this part of ISO/IEC 15938, the following terms and definitions apply.

3.3.1 Schema-related terminology

3.3.1.1
Attribute
A field of a **description tool** which is of simple type.

3.3.1.2
Base type
A **type** that serves as the root **type** of a derivation hierarchy for other **types**.

3.3.1.3
Datatype
A primitive reusable **type** employed by **Description Schemes** and **Descriptors**.

3.3.1.4
Derived type
A **type** that is defined in terms of extension or restriction of other **types**.

3.3.1.5
Description
An instantiation of one or more **description tools**.

3.3.1.6
Description Scheme
A **description tool** that describes entities or relationships pertaining to **multimedia content**. **Description Schemes** specify the structure and semantics of their components, which may be **Description Schemes**, **Descriptors**, or **datatypes**.

3.3.1.7
Description Tool
A **Description Scheme**, **Descriptor**, or **datatype**.

3.3.1.8
Descriptor
A **description tool** that describes a feature, attribute, or group of attributes of multimedia content.

3.3.1.9
Instantiation
Assignment of values to the fields (elements, attributes) of one or more **description tools**.

3.3.1.10
Element
A field of a **description tool** which is of complex type.

3.3.1.11
Schema
The set of related **description tools**, for example, those specified in ISO/IEC 15938.

3.3.1.12
Type
The format used for collection of letters, digits, and/or symbols, to depict values of an element or attribute of **description tool**. A **type** consists of a set of distinct values, a set of lexical representations, and a set of facets that characterize properties of the value space, individual values, or lexical items.

3.3.2   Content-related terminology

3.3.2.1
Abstraction
A secondary representation that is created from or is related to the **content**. For example, a **summary** of a **video** or a **model** of a **feature**.

3.3.2.2
Acquisition
The process of acquiring **audio** or **visual** data from a source.

3.3.2.3
Action
A semantically identifiable behaviour of an object or group of objects, for example, a soccer player kicking ball.

3.3.2.4
Agent
A person, organization, or group of persons.

3.3.2.5
Audio
Time-varying **data** or signal intended for listening or hearing. Also, related to the aural modality.

3.3.2.6
Audio-visual
content consisting of both **audio** and **video** data.

3.3.2.7
Automatic
Processing of **multimedia data**, **content**, or **metadata** by means of computer, hardware, or other software device.

3.3.2.8
Classification Scheme
A list of defined terms and their meanings.

3.3.2.9
Content
Multimedia content
A representation of the information contained in or related to **multimedia data** in a formalized manner suitable for interpretation by human means. **Content** refers to the **data** and the **metadata**.

3.3.2.10
Copyright
A right that establishes the ownership of **data**, **content**, or **metadata**.

3.3.2.11
Data
Essence
Multimedia Data
A representation of **multimedia** in a formalized manner suitable for communication, interpretation, or processing by automatic means.

3.3.2.12
Editing
The process of combining, extracting, and refining **multimedia data**.

3.3.2.13
Entity
Any concrete or abstract thing of interest related to the **multimedia content**.

3.3.2.14
Event
A noteworthy occurrence that happens at a point in time or during a temporal interval. Alternatively used as a change in state.

3.3.2.15
Feature
A distinctive characteristic of **multimedia content** that signifies something to a human observer, such as the "color" or "texture" of an image.

3.3.2.16
Filtering
A process for selecting multimedia content that satisfies certain criteria. This process may include ranking the content according to the extent that it satisfies the criteria.