

INTERNATIONAL  
STANDARD

ISO/IEC  
14496-12

Sixth edition  
2020-12

---

---

**Information technology — Coding of  
audio-visual objects —**

**Part 12:  
ISO base media file format**

*Technologies de l'information — Codage des objets audiovisuels —  
Partie 12: Format ISO de base pour les fichiers médias*



Reference number  
ISO/IEC 14496-12:2020(E)

© ISO/IEC 2020



**COPYRIGHT PROTECTED DOCUMENT**

© ISO/IEC 2020

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office  
CP 401 • Ch. de Blandonnet 8  
CH-1214 Vernier; Geneva  
Phone: +41 22 749 01 11  
Email: [copyright@iso.org](mailto:copyright@iso.org)  
Website: [www.iso.org](http://www.iso.org)

Published in Switzerland

# Contents

Page

<b>Foreword</b> .....	<b>ix</b>
<b>Introduction</b> .....	<b>x</b>
<b>1 Scope</b> .....	<b>1</b>
<b>2 Normative references</b> .....	<b>1</b>
<b>3 Terms, definitions and abbreviated terms</b> .....	<b>2</b>
3.1 Terms and definitions.....	2
3.2 Abbreviated terms.....	6
<b>4 Object-structured file organization</b> .....	<b>6</b>
4.1 File structure.....	6
4.2 Object structure.....	6
4.2.1 Object syntax conventions.....	6
4.2.2 Object definitions.....	7
4.3 File-type box.....	9
4.3.1 Definition.....	9
4.3.2 Syntax.....	9
4.3.3 Semantics.....	9
4.4 Extended type box.....	10
4.4.1 Definition.....	10
4.4.2 Syntax.....	10
4.4.3 Semantics.....	10
<b>5 Design considerations</b> .....	<b>10</b>
5.1 Usage.....	10
5.1.1 Multi-purpose.....	10
5.1.2 Interchange.....	11
5.1.3 Content creation.....	11
5.1.4 Preparation for streaming.....	11
5.1.5 Local presentation.....	12
5.1.6 Streamed presentation.....	12
5.2 Design principles.....	12
<b>6 ISO base media file organization</b> .....	<b>13</b>
6.1 Presentation structure.....	13
6.1.1 Referencing external data.....	13
6.1.2 Object structure.....	13
6.1.3 Meta data and media data.....	13
6.1.4 Track identifiers.....	13
6.2 Metadata structure (objects).....	14
6.2.1 Box.....	14
6.2.2 Data types and fields.....	14
6.2.3 Box order.....	15
6.2.4 URIs as type indicators.....	18
6.3 Brand identification.....	19
6.4 Time structure overview.....	19
<b>7 Streaming support</b> .....	<b>19</b>
7.1 Handling of streaming protocols.....	19
7.2 Protocol 'hint' tracks.....	20
7.3 Hint track format.....	20
<b>8 Box structures</b> .....	<b>22</b>
8.1 File structure and general boxes.....	22
8.1.1 Media data box.....	22
8.1.2 Free space box.....	22
8.1.3 Progressive download information box.....	23

8.1.4	Identified media data box .....	23
8.2	Movie structure .....	24
8.2.1	Movie box .....	24
8.2.2	Movie header box .....	24
8.3	Track structure .....	25
8.3.1	Track box .....	25
8.3.2	Track header box .....	26
8.3.3	Track reference box .....	29
8.3.4	Track group box .....	30
8.3.5	Track type box .....	32
8.4	Track media structure .....	32
8.4.1	Media box .....	32
8.4.2	Media header box .....	32
8.4.3	Handler reference box .....	33
8.4.4	Media information box .....	34
8.4.5	Media information header boxes .....	34
8.4.6	Extended language tag .....	35
8.5	Sample tables .....	36
8.5.1	Sample table box .....	36
8.5.2	Sample description box .....	36
8.5.3	Degradation priority box .....	38
8.5.4	Sample scale box .....	39
8.6	Track time structures .....	39
8.6.1	Time to sample boxes .....	39
8.6.2	Sync sample box .....	44
8.6.3	Shadow sync sample box .....	44
8.6.4	Independent and disposable samples box .....	45
8.6.5	Edit box .....	47
8.6.6	Edit list box .....	47
8.7	Track data layout structures .....	49
8.7.1	Data information box .....	49
8.7.2	Data reference box .....	50
8.7.3	Sample size boxes .....	51
8.7.4	Sample to chunk box .....	53
8.7.5	Chunk offset box .....	53
8.7.6	Padding bits box .....	54
8.7.7	Sub-sample information box .....	55
8.7.8	Sample auxiliary information sizes box .....	57
8.7.9	Sample auxiliary information offsets box .....	58
8.8	Movie fragments .....	59
8.8.1	Movie extends box .....	59
8.8.2	Movie extends header box .....	60
8.8.3	Track extends box .....	60
8.8.4	Movie fragment box .....	61
8.8.5	Movie fragment header box .....	62
8.8.6	Track fragment box .....	62
8.8.7	Track fragment header box .....	63
8.8.8	Track fragment run box .....	64
8.8.9	Movie fragment random access box .....	65
8.8.10	Track fragment random access box .....	66
8.8.11	Movie fragment random access offset box .....	67
8.8.12	Track fragment decode time box .....	67
8.8.13	Level assignment box .....	68
8.8.14	Sample auxiliary information in movie fragments .....	70
8.8.15	Track Extension Properties box .....	70
8.8.16	Alternative startup sequence properties box .....	71
8.8.17	Metadata and user data in movie fragments .....	72
8.9	Sample group structures .....	72

8.9.1	Overview	72
8.9.2	Sample to group box	73
8.9.3	Sample group description box	74
8.9.4	Representation of group structures in movie fragments	76
8.9.5	Compact sample to group box	77
8.10	User data	79
8.10.1	User data box	79
8.10.2	Copyright box	79
8.10.3	Track selection box	80
8.10.4	Track kind	81
8.11	Metadata support	82
8.11.1	MetaBox	82
8.11.2	XML boxes	83
8.11.3	Item location box	84
8.11.4	Primary item box	86
8.11.5	Item protection box	87
8.11.6	Item information box	87
8.11.7	Additional metadata container box	90
8.11.8	Metabox Relation box	90
8.11.9	URL forms for MetaBoxes	90
8.11.10	Static metadata	90
8.11.11	Item data box	91
8.11.12	Item reference box	92
8.11.13	Auxiliary video metadata	93
8.11.14	Item properties box	93
8.11.15	Brand item property	94
8.12	Support for protected streams	95
8.12.1	Overview	95
8.12.2	Protection scheme information box	96
8.12.3	Original format box	97
8.12.4	IPMPInfoBox	97
8.12.5	IPMP control box	97
8.12.6	Scheme type box	97
8.12.7	Scheme information box	98
8.12.8	Scramble Scheme Information Box	98
8.13	File delivery format support	99
8.13.1	Overview	99
8.13.2	FD item information box	99
8.13.3	File partition box	100
8.13.4	FEC reservoir box	101
8.13.5	FD session group box	102
8.13.6	Group ID to name box	103
8.13.7	File reservoir box	103
8.14	Sub tracks	104
8.14.1	Overview	104
8.14.2	Backward compatibility	104
8.14.3	Sub track box	105
8.14.4	Sub track information box	105
8.14.5	Sub track definition box	106
8.14.6	Sub track sample group box	106
8.15	Post-decoder requirements on media	107
8.15.1	General	107
8.15.2	Restricted sample entry transformation	107
8.15.3	Restricted scheme information box	108
8.15.4	Scheme for stereoscopic video arrangements	108
8.15.5	Compatible scheme type box	110
8.16	Segments	111
8.16.1	Overview	111

8.16.2	Segment type box	111
8.16.3	Segment index box	111
8.16.4	Subsegment index box	115
8.16.5	Producer reference time box	116
8.17	Support for incomplete tracks	118
8.17.1	General	118
8.17.2	Transformation	119
8.17.3	Complete track information box	119
8.18	Entity grouping	119
8.18.1	General	119
8.18.2	Groups list box	120
8.18.3	Entity to group box	120
8.19	Compressed boxes	121
8.19.1	Overview and processing	121
8.19.2	Processing model	121
8.19.3	General syntax	123
8.19.4	General semantics	123
8.19.5	Original file-type box	123
8.19.6	Compressed movie box	123
8.19.7	Compressed movie fragment box	124
8.19.8	Compressed segment index box	124
8.19.9	Compressed subsegment index box	125
<b>9</b>	<b>Hint track formats</b>	<b>125</b>
9.1	RTP and SRTP hint track format	125
9.1.1	Overview	125
9.1.2	Sample description format	126
9.1.3	Sample format	127
9.1.4	SDP information	129
9.1.5	Statistical information	130
9.2	ALC/LCT and FLUTE hint track format	131
9.2.1	Overview	131
9.2.2	Design principles	131
9.2.3	Sample description format	133
9.2.4	Sample format	133
9.3	MPEG-2 transport hint track format	136
9.3.1	Overview	136
9.3.2	Design principles	137
9.3.3	Sample description format	138
9.3.4	Sample format	140
9.3.5	Protected MPEG 2 transport stream hint track	142
9.4	RTP, RTCP, SRTP and SRTCP reception hint tracks	143
9.4.1	RTP reception hint track	143
9.4.2	RTCP reception hint track	146
9.4.3	SRTP reception hint track	147
9.4.4	SRTCP reception hint tracks	149
9.4.5	Protected RTP reception hint track	150
9.4.6	Recording procedure	150
9.4.7	Parsing procedure	150
<b>10</b>	<b>Sample groups</b>	<b>150</b>
10.1	Random access recovery points	150
10.1.1	Definition	150
10.1.2	Syntax	151
10.1.3	Semantics	151
10.2	Rate share groups	151
10.2.1	Overview	151
10.2.2	Rate share sample group entry	152
10.2.3	Relationship between tracks	153

10.2.4	Bitrate allocation .....	154
10.3	Alternative startup sequences .....	154
10.3.1	Definition .....	154
10.3.2	Syntax .....	155
10.3.3	Semantics .....	155
10.3.4	Examples .....	155
10.4	Random access point (RAP) sample group .....	157
10.4.1	Definition .....	157
10.4.2	Syntax .....	157
10.4.3	Semantics .....	157
10.5	Temporal level sample group .....	157
10.5.1	Definition .....	157
10.5.2	Syntax .....	158
10.5.3	Semantics .....	158
10.6	Stream access point sample group .....	158
10.6.1	Definition .....	158
10.6.2	Syntax .....	158
10.6.3	Semantics .....	158
10.7	Sample-to-item sample group .....	159
10.7.1	Definition .....	159
10.7.2	Syntax .....	159
10.7.3	Semantics .....	159
10.8	Dependent random access point (DRAP) sample group .....	159
10.8.1	Definition .....	159
10.8.2	Syntax .....	160
10.8.3	Semantics .....	160
<b>11</b>	<b>Extensibility .....</b>	<b>160</b>
11.1	Objects .....	160
11.2	Storage formats .....	161
11.3	Derived file formats .....	161
<b>12</b>	<b>Media-specific definitions .....</b>	<b>162</b>
12.1	Video media .....	162
12.1.1	Media handler .....	162
12.1.2	Video media header .....	162
12.1.3	Sample entry .....	162
12.1.4	Pixel aspect ratio and clean aperture .....	163
12.1.5	Colour information .....	165
12.1.6	Content light level .....	166
12.1.7	Mastering display colour volume .....	166
12.1.8	Content colour volume .....	166
12.2	Audio media .....	167
12.2.1	Media handler .....	167
12.2.2	Sound media header .....	167
12.2.3	Sample entry .....	168
12.2.4	Channel layout .....	170
12.2.5	Downmix instructions .....	172
12.2.6	DRC information .....	175
12.2.7	Audio stream loudness .....	176
12.3	Metadata media .....	178
12.3.1	Media handler .....	178
12.3.2	Media header .....	178
12.3.3	Sample entry .....	178
12.4	Hint media .....	180
12.4.1	Media handler .....	180
12.4.2	Hint media header .....	180
12.4.3	Sample entry .....	180
12.5	Text media .....	181

12.5.1	Media handler .....	181
12.5.2	Media header .....	181
12.5.3	Sample entry .....	181
12.6	Subtitle media .....	181
12.6.1	Media handler .....	181
12.6.2	Subtitle media header .....	181
12.6.3	Sample entry .....	182
12.7	Font media .....	183
12.7.1	Media handler .....	183
12.7.2	Media header .....	183
12.7.3	Sample entry .....	183
12.8	Transformed media .....	183
12.8.1	General .....	183
12.8.2	Multiple transformations for a single transformed media track .....	183
12.8.3	Determining the untransformed sample entry type .....	184
12.8.4	The 'codecs' MIME parameter for a transformed media track .....	184
12.9	Multiplexed timed metadata tracks .....	184
12.9.1	General .....	184
12.9.2	Overall design .....	184
12.9.3	Sample format .....	185
12.9.4	Sample entry format .....	185
12.9.5	Defined formats .....	188
<b>Annex A (informative) Overview of the file format .....</b>		<b>190</b>
<b>Annex B (informative) Guidance on deriving from this document .....</b>		<b>194</b>
<b>Annex C (normative) Fragment identifiers for ISO base media resources .....</b>		<b>203</b>
<b>Annex D (informative) Management of extension code points .....</b>		<b>204</b>
<b>Annex E (normative) File format brands .....</b>		<b>205</b>
<b>Annex F (normative) MIME type registration of segments .....</b>		<b>217</b>
<b>Annex G (informative) URI-labelled metadata forms .....</b>		<b>218</b>
<b>Annex H (informative) Processing of RTP streams and reception hint tracks .....</b>		<b>220</b>
<b>Annex I (normative) Stream access points .....</b>		<b>235</b>
<b>Annex J (informative) Segment index examples .....</b>		<b>238</b>
<b>Annex K (normative) Use of IETF RFC 6381 for ISOBMFF files .....</b>		<b>241</b>
<b>Bibliography .....</b>		<b>244</b>



## 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.

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](http://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](http://www.iso.org/patents)) or the IEC list of patent declarations received (see [patents.iec.ch](http://patents.iec.ch)).

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 <https://www.iso.org/foreword-supplementary-information.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*.

This sixth edition cancels and replaces the fifth edition (ISO/IEC 14496-12:2015), which has been technically revised. It also incorporates Amendments ISO/IEC 14496-12:2015/Amd.1:2017 and ISO/IEC 14496-12:2015/Amd.2:2018.

The main changes compared to the previous edition are as follows:

- incorporation of all changes provided in Amendments 1 and 2 and unpublished draft COR.1 for the fifth edition and draft Amendments 1 through 4 for this sixth edition.
- editorial changes to align the document with the drafting rules in ISO/IEC Directives Part 2.

A list of all parts in the ISO/IEC 14496 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 <https://www.iso.org/members.html>.

## Introduction

The ISO base media file format is designed to contain timed media information for a presentation in a flexible, extensible format that facilitates interchange, management, editing, and presentation of the media. This presentation may be 'local' to the system containing the presentation, or may be via a network or other stream delivery mechanism.

The file structure is object-oriented; a file can be decomposed into constituent objects very simply, and the structure of the objects inferred directly from their type.

The file format is designed to be independent of any particular network protocol while enabling efficient support for them in general.

The ISO base media file format is a base format for media file formats.

The International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC) draw attention to the fact that it is claimed that compliance with this document may involve the use of patents.

ISO and IEC take no position concerning the evidence, validity and scope of these patent rights

The holders of these patent rights have assured ISO and IEC that they are willing to negotiate licences under reasonable and non-discriminatory terms and conditions with applicants throughout the world. In this respect, the statements of the holders of these patent rights are registered with ISO and IEC. Information may be obtained from the patent database available at [www.iso.org/patents](http://www.iso.org/patents).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights other than those in the patent database. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

# Information technology — Coding of audio-visual objects —

## Part 12: ISO base media file format

### 1 Scope

This document specifies the ISO base media file format, which is a general format forming the basis for a number of other more specific file formats. This format contains the timing, structure, and media information for timed sequences of media data, such as audio-visual presentations.

### 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 639-2, *Codes for the representation of names of languages — Part 2: Alpha-3 code*

ITU-T X.667 | ISO/IEC 9834-8, *Information technology — Procedures for the operation of object identifier registration authorities — Part 8: Generation of universally unique identifiers (UUIDs) and their use in object identifiers*

ISO/IEC 10646, *Information technology — Universal Coded Character Set (UCS)*

ISO/IEC 13818-2:2013, *Information technology — Generic coding of moving pictures and associated audio information — Part 2: Video*

ISO/IEC 14496-1, *Information technology — Coding of audio-visual objects — Part 1: Systems*

ISO/IEC 14496-10:2014, *Information technology — Coding of audio-visual objects — Part 10: Advanced Video Coding*

ISO 15076-1, *Image technology colour management — Architecture, profile format and data structure — Part 1: Based on ICC.1:2010*

ISO/IEC 15938-1, *Information technology — Multimedia content description interface — Part 1: Systems*

ISO/IEC 23001-1, *Information technology — MPEG systems technologies — Part 1: Binary MPEG format for XML*

ISO/IEC 23002-3, *Information technology — MPEG video technologies — Part 3: Representation of auxiliary video and supplemental information*

ISO/IEC 23003-4, *Information technology — MPEG audio technologies — Part 4: Dynamic range control*

ITU-T H.265 | ISO/IEC 23008-2, *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 2: High efficiency video coding*

ISO/IEC 23091-2, *Information technology — Coding-independent code points — Part 2: Video*

ISO/IEC 23091-3, *Information technology — Coding-independent code points — Part 3: Audio*

IETF RFC 1951, *DEFLATE Compressed Data Format Specification version 1.3*

## ISO/IEC 14496-12:2020(E)

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

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

IETF RFC 3629, *UTF-8, a transformation format of ISO 10646*

IETF RFC 3711:2004, *The Secure Real-time Transport Protocol (SRTP)*

IETF RFC 5052, *Forward Error Correction (FEC) Building Block*

IETF RFC 5905, *Network Time Protocol Version 4: Protocol and Algorithms Specification*

ITU-R TF.460-6:2002, *Standard-frequency and time-signal emissions*

ITU-R BS.1770-4, *Algorithms to measure audio programme loudness and true-peak audio level*

IETF BCP 47, *Tags for Identifying Languages*

IETF RFC 4122, *A Universally Unique Identifier (UUID) URN Namespace*

IETF RFC 3061, *A URN Namespace of Object Identifiers*

W3C Recommendation, *Extensible Markup Language (XML) 1.0 (Fifth Edition)*, 26 November 2008, <https://www.w3.org/TR/2008/REC-xml-20081126/>