

FINAL
DRAFT

INTERNATIONAL
STANDARD

ISO/IEC
FDIS
14496-12

ISO/IEC JTC 1/SC 29

Secretariat: JISC

Voting begins on:
2020-06-19

Voting terminates on:
2020-08-14

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*

iTeh STANDARD PREVIEW
(standards.iteh.ai)
Full standard:
<https://standards.iteh.ai/catalog/standards/iso-iec-fdis/14496-12>
21bd-44ec-b9b2-7ea80caaf0ad0/iso-iec-fdis/14496-12

RECIPIENTS OF THIS DRAFT ARE INVITED TO SUBMIT, WITH THEIR COMMENTS, NOTIFICATION OF ANY RELEVANT PATENT RIGHTS OF WHICH THEY ARE AWARE AND TO PROVIDE SUPPORTING DOCUMENTATION.

IN ADDITION TO THEIR EVALUATION AS BEING ACCEPTABLE FOR INDUSTRIAL, TECHNOLOGICAL, COMMERCIAL AND USER PURPOSES, DRAFT INTERNATIONAL STANDARDS MAY ON OCCASION HAVE TO BE CONSIDERED IN THE LIGHT OF THEIR POTENTIAL TO BECOME STANDARDS TO WHICH REFERENCE MAY BE MADE IN NATIONAL REGULATIONS.



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

© ISO/IEC 2020

iTeh STANDARD PREVIEW
(standards.iteh.ai)
Full standard:
<https://standards.iteh.ai/catalog/standards/sist/1d84f475-21bd-44ec-b9b2-7ea80caff0ad0/iso-iec-fdis-14496-12>



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
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword	ix
Introduction	x
1 Scope	1
2 Normative references	1
3 Terms, definitions and abbreviated terms	2
3.1 Terms and definitions.....	2
3.2 Abbreviated terms.....	6
4 Object-structured file organization	6
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.....	8
4.3.1 Definition.....	8
4.3.2 Syntax.....	9
4.3.3 Semantics.....	9
4.4 Extended type box.....	9
4.4.1 Definition.....	9
4.4.2 Syntax.....	10
4.4.3 Semantics.....	10
5 Design considerations	10
5.1 Usage.....	10
5.1.1 Multi-purpose.....	10
5.1.2 Interchange.....	10
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
6 ISO base media file organization	13
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
7 Streaming support	19
7.1 Handling of streaming protocols.....	19
7.2 Protocol 'hint' tracks.....	20
7.3 Hint track format.....	21
8 Box structures	22
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	28
8.3.4	Track group box	30
8.3.5	Track type box	31
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	34
8.5	Sample tables	35
8.5.1	Sample table box	35
8.5.2	Sample description box	36
8.5.3	Degradation priority box	37
8.5.4	Sample scale box	38
8.6	Track time structures	38
8.6.1	Time to sample boxes	38
8.6.2	Sync sample box	43
8.6.3	Shadow sync sample box	43
8.6.4	Independent and disposable samples box	44
8.6.5	Edit box	46
8.6.6	Edit list box	46
8.7	Track data layout structures	49
8.7.1	Data information box	49
8.7.2	Data reference box	49
8.7.3	Sample size boxes	51
8.7.4	Sample to chunk box	52
8.7.5	Chunk offset box	53
8.7.6	Padding bits box	53
8.7.7	Sub-sample information box	54
8.7.8	Sample auxiliary information sizes box	55
8.7.9	Sample auxiliary information offsets box	57
8.8	Movie fragments	58
8.8.1	Movie extends box	58
8.8.2	Movie extends header box	59
8.8.3	Track extends box	59
8.8.4	Movie fragment box	60
8.8.5	Movie fragment header box	61
8.8.6	Track fragment box	61
8.8.7	Track fragment header box	61
8.8.8	Track fragment run box	63
8.8.9	Movie fragment random access box	64
8.8.10	Track fragment random access box	64
8.8.11	Movie fragment random access offset box	66
8.8.12	Track fragment decode time box	66
8.8.13	Level assignment box	67
8.8.14	Sample auxiliary information in movie fragments	69
8.8.15	Track Extension Properties box	69
8.8.16	Alternative startup sequence properties box	69
8.8.17	Metadata and user data in movie fragments	70
8.9	Sample group structures	71

8.9.1	Overview	71
8.9.2	Sample to group box	71
8.9.3	Sample group description box	73
8.9.4	Representation of group structures in movie fragments	74
8.9.5	Compact sample to group box	75
8.10	User data	77
8.10.1	User data box	77
8.10.2	Copyright box	78
8.10.3	Track selection box	78
8.10.4	Track kind	80
8.11	Metadata support	80
8.11.1	MetaBox	80
8.11.2	XML boxes	81
8.11.3	Item location box	82
8.11.4	Primary item box	85
8.11.5	Item protection box	85
8.11.6	Item information box	86
8.11.7	Additional metadata container box	88
8.11.8	Metabox Relation box	88
8.11.9	URL forms for MetaBoxes	88
8.11.10	Static metadata	89
8.11.11	Item data box	90
8.11.12	Item reference box	90
8.11.13	Auxiliary video metadata	91
8.11.14	Item properties box	91
8.11.15	Brand item property	93
8.12	Support for protected streams	93
8.12.1	Overview	93
8.12.2	Protection scheme information box	94
8.12.3	Original format box	95
8.12.4	IPMPInfoBox	95
8.12.5	IPMP control box	95
8.12.6	Scheme type box	96
8.12.7	Scheme information box	96
8.12.8	Scramble Scheme Information Box	96
8.13	File delivery format support	97
8.13.1	Overview	97
8.13.2	FD item information box	97
8.13.3	File partition box	98
8.13.4	FEC reservoir box	99
8.13.5	FD session group box	100
8.13.6	Group ID to name box	101
8.13.7	File reservoir box	102
8.14	Sub tracks	102
8.14.1	Overview	102
8.14.2	Backward compatibility	103
8.14.3	Sub track box	103
8.14.4	Sub track information box	103
8.14.5	Sub track definition box	104
8.14.6	Sub track sample group box	105
8.15	Post-decoder requirements on media	105
8.15.1	General	105
8.15.2	Restricted sample entry transformation	105
8.15.3	Restricted scheme information box	106
8.15.4	Scheme for stereoscopic video arrangements	106
8.15.5	Compatible scheme type box	108
8.16	Segments	109
8.16.1	Overview	109

8.16.2	Segment type box	109
8.16.3	Segment index box	110
8.16.4	Subsegment index box	113
8.16.5	Producer reference time box	115
8.17	Support for incomplete tracks	116
8.17.1	General	116
8.17.2	Transformation	117
8.17.3	Complete track information box	117
8.18	Entity grouping	117
8.18.1	General	117
8.18.2	Groups list box	118
8.18.3	Entity to group box	118
8.19	Compressed boxes	119
8.19.1	Overview and processing	119
8.19.2	Processing model	119
8.19.3	General syntax	121
8.19.4	General semantics	121
8.19.5	Original file-type box	121
8.19.6	Compressed movie box	121
8.19.7	Compressed movie fragment box	122
8.19.8	Compressed segment index box	122
8.19.9	Compressed subsegment index box	123
9	Hint track formats	123
9.1	RTP and SRTP hint track format	123
9.1.1	Overview	123
9.1.2	Sample description format	124
9.1.3	Sample format	125
9.1.4	SDP information	127
9.1.5	Statistical information	128
9.2	ALC/LCT and FLUTE hint track format	129
9.2.1	Overview	129
9.2.2	Design principles	129
9.2.3	Sample description format	131
9.2.4	Sample format	131
9.3	MPEG-2 transport hint track format	134
9.3.1	Overview	134
9.3.2	Design principles	135
9.3.3	Sample description format	136
9.3.4	Sample format	138
9.3.5	Protected MPEG 2 transport stream hint track	140
9.4	RTP, RTCP, SRTP and SRTCP reception hint tracks	141
9.4.1	RTP reception hint track	141
9.4.2	RTCP reception hint track	144
9.4.3	SRTP reception hint track	145
9.4.4	SRTCP reception hint tracks	147
9.4.5	Protected RTP reception hint track	148
9.4.6	Recording procedure	148
9.4.7	Parsing procedure	148
10	Sample groups	148
10.1	Random access recovery points	148
10.1.1	Definition	148
10.1.2	Syntax	149
10.1.3	Semantics	149
10.2	Rate share groups	149
10.2.1	Overview	149
10.2.2	Rate share sample group entry	150
10.2.3	Relationship between tracks	151

10.2.4	Bitrate allocation	152
10.3	Alternative startup sequences	152
10.3.1	Definition	152
10.3.2	Syntax	153
10.3.3	Semantics	153
10.3.4	Examples	153
10.4	Random access point (RAP) sample group	155
10.4.1	Definition	155
10.4.2	Syntax	155
10.4.3	Semantics	155
10.5	Temporal level sample group	155
10.5.1	Definition	155
10.5.2	Syntax	156
10.5.3	Semantics	156
10.6	Stream access point sample group	156
10.6.1	Definition	156
10.6.2	Syntax	156
10.6.3	Semantics	156
10.7	Sample-to-item sample group	157
10.7.1	Definition	157
10.7.2	Syntax	157
10.7.3	Semantics	157
10.8	Dependent random access point (DRAP) sample group	157
10.8.1	Definition	157
10.8.2	Syntax	158
10.8.3	Semantics	158
11	Extensibility	158
11.1	Objects	158
11.2	Storage formats	159
11.3	Derived file formats	159
12	Media-specific definitions	159
12.1	Video media	159
12.1.1	Media handler	159
12.1.2	Video media header	160
12.1.3	Sample entry	160
12.1.4	Pixel aspect ratio and clean aperture	161
12.1.5	Colour information	163
12.1.6	Content light level	163
12.1.7	Mastering display colour volume	164
12.1.8	Content colour volume	164
12.2	Audio media	165
12.2.1	Media handler	165
12.2.2	Sound media header	165
12.2.3	Sample entry	165
12.2.4	Channel layout	167
12.2.5	Downmix instructions	169
12.2.6	DRC information	173
12.2.7	Audio stream loudness	173
12.3	Metadata media	176
12.3.1	Media handler	176
12.3.2	Media header	176
12.3.3	Sample entry	176
12.4	Hint media	177
12.4.1	Media handler	177
12.4.2	Hint media header	177
12.4.3	Sample entry	178
12.5	Text media	178

12.5.1	Media handler	178
12.5.2	Media header	178
12.5.3	Sample entry	178
12.6	Subtitle media	179
12.6.1	Media handler	179
12.6.2	Subtitle media header	179
12.6.3	Sample entry	179
12.7	Font media	180
12.7.1	Media handler	180
12.7.2	Media header	180
12.7.3	Sample entry	180
12.8	Transformed media	180
12.8.1	General	180
12.8.2	Multiple transformations for a single transformed media track	181
12.8.3	Determining the untransformed sample entry type	181
12.8.4	The 'codecs' MIME parameter for a transformed media track	182
12.9	Multiplexed timed metadata tracks	182
12.9.1	General	182
12.9.2	Overall design	182
12.9.3	Sample format	182
12.9.4	Sample entry format	182
12.9.5	Defined formats	185
Annex A (informative) Overview of the file format		187
Annex B (informative) Guidance on deriving from this document		191
Annex C (normative) Fragment identifiers for ISO base media resources		200
Annex D (informative) Management of extension code points		201
Annex E (normative) File format brands		202
Annex F (normative) MIME type registration of segments		214
Annex G (informative) URI-labelled metadata forms		215
Annex H (informative) Processing of RTP streams and reception hint tracks		217
Annex I (normative) Stream access points		232
Annex J (informative) Segment index examples		235
Annex K (normative) Use of IETF RFC 6381 for ISOBMFF files		238
Bibliography		241

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).

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 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.

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*

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 BC 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/>

3 Terms, definitions and abbreviated terms

3.1 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:

- ISO Online browsing platform: available at <https://www.iso.org/obp>
- IEC Electropedia: available at <http://www.electropedia.org/>

3.1.1

box

object-oriented building block defined by a unique type identifier and length

Note 1 to entry: Called 'atom' in some specifications, including the first definition of MP4.

3.1.2

chunk

contiguous set of samples for one track

3.1.3

clean aperture

part of a decoded video image from which undesirable pixels introduced for coding purposes such as having integer number of coding blocks have been removed for presentation

3.1.4

container box

box whose sole purpose is to contain and group a set of related boxes

Note 1 to entry: Container boxes are normally not derived from FullBox.

3.1.5

file level

byte position in an ISO base media file not contained in a Box structure

3.1.6**full aperture**

decoded video image as output by the decoder which may contain undesirable pixels for presentation

3.1.7**hint track**

special track which does not contain media data, but instead contains instructions for packaging one or more tracks into a streaming channel

3.1.8**hinter**

tool that is run on a file containing only media, to add one or more hint tracks to the file and so facilitate streaming

3.1.9**index file**

ISO base media file containing only `SegmentIndexBox`

3.1.10**ISO base media file**

file conforming to the file format described in this document (either a movie file, a metadata file, a segment file or an index file)

3.1.11**item**

data which does not require timed processing, as opposed to sample data

3.1.12**leading sample**

sample associated with a random access point (RAP) that precedes the RAP in composition order and immediately follows the RAP or another leading sample in decoding order, and which possibly cannot be correctly decoded when decoding starts from the RAP

3.1.13**leaf subsegment**

subsegment that does not contain any indexing information that would enable its further division into subsegments

3.1.14**mod**

modulo operator: $(x \text{ mod } y) = x - y \text{ floor } (x/y)$

3.1.15**media data box**

box which can hold the actual media data for a presentation ('`mdat`')

3.1.16**metadata file**

ISO base media file containing a top-level `MetaBox`

Note 1 to entry: A Movie File may also be a Metadata File, and vice-versa.

3.1.17**movie box**

container box whose sub-boxes define the metadata for a presentation ('`moov`')

3.1.18**movie file**

ISO base media file containing a `MovieBox`