International Standard

ISO/IEC 14496-12

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

Eighth edition
2026-04

Reference number
ISO/IEC 14496-12:2026(en)

© ISO/IEC 2026

COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2026

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

© ISO/IEC 2026 – All rights reserved

© ISO/IEC 2026 – All rights reserved

© ISO/IEC 2026 – All rights reserved

© ISO/IEC 2026 – All rights reserved

© ISO/IEC 2026 – All rights reserved

© ISO/IEC 2026 – All rights reserved

© ISO/IEC 2026 – All rights reserved

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 or www.iec.ch/members_experts/refdocs).

ISO and IEC draw attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO and IEC take 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 and IEC had 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 and https://patents.iec.ch. ISO and IEC 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. In the IEC, see www.iec.ch/understanding-standards.

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

This eigth edition cancels and replaces the seventh edition (ISO/IEC 14496-12:2022), which has been technically revised.

The main changes are as follows:

— specification of essential sample groups

— addition of the essential descriptions hierarchy sample group

— addition of the preselection entity group and related boxes

— addition of the extended dependent random access point (EDRAP) sample group

— specification of the sample-packed restricted video track

— addition of the associated external stream track reference

— relaxing the presence requirements of the `HandlerBox` within a `MetaBox`

— addition of the handler property for items.

A list of all parts in the ISO/IEC 14496 series can be found on the ISO and IEC websites.

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 and www.iec.ch/national-committees.

© ISO/IEC 2026 – All rights reserved

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.

Structure of this document

Clause 4 defines ISO base media files; files that are built from boxes.

Clause 5 defines typed ISO base media files, i.e. ISO base media files with a `FileTypeBox`.

Clause 6 provides common specifications to movie files, item files and segment files.

Clause 7 gives the core concepts and data-types for time-based presentations, called 'movies' in this document.

Clause 8 defines the boxes used by time-based presentations, and other formats.

Clause 9 defines the hint track formats used to support some streaming protocols.

Clause 10 builds on the concept of sample groups as defined in subclause 8.9 and defines some sample groups.

Clause 11 defines how to base a file format on this document.

Clause 12 builds on the general concepts of tracks as defined in Clause 8, and defines track formats for various general types of media (video, sound, etc.).

Clause 13 specifies different types of transformed media tracks.

Annex A provides an informative introduction to time-based presentations, which may be of assistance to first-time readers and implementers.

Annex B provides guidance on writing derived specifications.

Annex C provides the syntax for uniform resource identifier (URI) fragments.

Annex D documents how identifier values defined externally to this document are managed.

Annex E defines brands that may be used to identify the conformance and reader requirements to the structures defined in this document for time-based presentations.

Annex F contains the formal IANA registration of segments.

Annex G defines some forms used for labelling metadata with uniform resource identifier (URI) labels.

Annex H provides an overview of the use of hint tracks for RTP streams and RTP stream reception.

Annex I contains the formal definitions of the types of stream access points in timed media streams.

Annex J contains examples of the use of the `SegmentIndexBox` defined in 8.14.3.

Annex K defines the MIME parameters that may be used to annotate MIME types for time-based presentations based on Clause 6.

© ISO/IEC 2026 – All rights reserved

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, *Code for individual languages and language groups*

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

ISO/IEC 14496-34, *Information Information technology — Coding of audio-visual objects — Part 34: Syntactic description language*

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 23001-14, *Information technology — MPEG systems technologies — Part 14: Partial file format*

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*

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*

© ISO/IEC 2026 – All rights reserved

ITU-T T.35, *Procedure for the allocation of ITU-T defined codes for non-standard facilities*

IETF BCP 47, *Tags for Identifying Languages*

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 3061, *A URN Namespace of Object Identifiers*

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

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

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

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

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

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

© ISO/IEC 2026 – All rights reserved

Bibliography

[1] ISO/IEC 8859-1, *Information technology — 8-bit single-byte coded graphic character sets — Part 1: Latin alphabet No. 1*

[2] ISO/IEC 13818-1, *Information technology — Generic coding of moving pictures and associated audio information — Part 1: Systems*

[3] ISO/IEC 14496-14, *Information technology — Coding of audio-visual objects — Part 14: MP4 file format*

[4] ISO/IEC 14496-15, *Information technology — Coding of audio-visual objects — Part 15: Carriage of network abstraction layer (NAL) unit structured video in the ISO base media file format*

[5] ISO/IEC 15444-3, *Information technology — JPEG 2000 image coding system: Motion JPEG 2000 — Part 3:*

[6] ISO/IEC 21000-17:2006, *Information technology — Multimedia framework (MPEG-21) — Part 17: Fragment Identification of MPEG Resources*

[7] ISO/IEC 23000-11, *Information technology — Multimedia application format (MPEG-A) — Part 11: Stereoscopic video application format*

[8] ISO/IEC 23008-3, *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 3: 3D audio*

[9] ISO/IEC 23008-12, *Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 12: Image File Format*

[10] ISO/IEC 23009-1:2022, *Information technology — Dynamic adaptive streaming over HTTP (DASH) — Part 1: Media presentation description and segment formats*

[11] Rec. ITU-R BS.1771-1, *Requirements for loudness and true-peak indicating meters*

[12] IETF RFC 1864, *The Content-MD5 Header Field, October 1995*

[13] IETF RFC 2616, *Hypertext Transfer Protocol — HTTP/1.1, June 1999*

[14] IETF RFC 3450, *Asynchronous Layered Coding (ALC) Protocol Instantiation, December 2002*

[15] IETF RFC 3451, *Layered Coding Transport (LCT) Building Block, December 2002*

[16] IETF RFC 3550, *RTP: A Transport Protocol for Real-Time Applications, July 2003*

[17] IETF RFC 3551, *RTP Profile for Audio and Video Conferences with Minimal Control, July 2003*

[18] IETF RFC 3695, *Compact Forward Error Correction (FEC) Schemes, February 2004*

[19] IETF RFC 3926, *FLUTE - File Delivery over Unidirectional Transport, October 2004*

[20] IETF RFC 3986, *Uniform Resource Identifier (URI): Generic Syntax, January 2005*

[21] IETF RFC 3987, *Internationalized Resource Identifiers (IRIs), January 2005*

[22] IETF RFC 4771, *Integrity Transform Carrying Roll-Over Counter for the Secure Real-time Transport Protocol (SRTP), January 2007*

[23] IETF RFC 5117, *RTP Topologies, January 2008*

[24] IETF RFC 5119, *A Uniform Resource Name (URN) Namespace for the Society of Motion Picture and Television Engineers (SMPTE), February 2008*

[25] IETF RFC 6381, *The 'Codecs' and 'Profiles' Parameters for "Bucket" Media Types, August 2011*

© ISO/IEC 2026 – All rights reserved

[26] ICC.1:2001-04, *File format for color profiles*, International Color Consortium

[27] The Unicode Standard. https://www.unicode.org/versions/latest/

[28] 3GPP TS 26.244, *3GPP file format (3GP)*

[29] 3GPP TS 26.346, *Multimedia Broadcast/Multicast Service (MBMS); Protocols and codecs*

[30] SMPTE RP 177, *Derivation of Basic Television Color Equations; Society of Motion Picture and Television Engineers (SMPTE), 1993*

[31] SMPTE ST 336:2017, *Data Encoding Protocol Using Key-Length-Value*

[32] EBU – Tech 3341, *Loudness Metering: EBU mode metering to supplement loudness normalization in accordance with EBU R128*

[33] EBU – Tech 3342, *Loudness Range: A measure to supplement loudness normalisation in accordance with EBU R 128*

[34] Doc A.T.S.C. A/52:2012, *ATSC Standard: Digital Audio Compression (AC-3, E-AC-3)*

[35] *Consumer Technology Association CTA-861-G*, A DTV Profile for Uncompressed High Speed Digital Interface

[36] 3GPP TS 26.071, *Mandatory speech CODEC speech processing functions; AMR speech Codec; General description*

[37] Rec. ITU-T H.263, *Video coding for low bit rate communication*

[38] ISO/IEC 11172-3, *Information technology — Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s — Part 3: Audio*

© ISO/IEC 2026 – All rights reserved