
**Information technology — Open Document
Architecture (ODA) and interchange format:
Audio content architectures**

*Technologies de l'information — Architecture de document ouverte (ODA)
et format de transfert: Architectures du système audio*



Reference number
ISO/IEC 8613-9:1996(E)

AVA **Documents**
Document Sharing Hub

CONTENTS

	<i>Page</i>
1 Scope	1
2 Normative references	2
2.1 Identical Recommendations International Standards	2
2.2 Paired Recommendations International Standards equivalent in technical content	2
2.3 Additional references	2
3 Definitions	3
4 Abbreviations	3
5 Conventions	3
6 General principles	3
6.1 Content architectures	3
6.1.1 Formatted processable content architecture class	4
6.2 Content	4
6.3 Presentation attributes	4
6.4 Content portion attributes	4
6.5 Coding of content information	4
7 Audio positioning	4
7.1 Basic concepts	4
7.1.1 Measurement units	4
7.2 Positioning of elements of sampled sound in a basic layout time frame	5
7.3 Reference recording level	5
8 Audio presentation	5
8.1 Attenuation and temporal envelope	5
8.2 Frequency envelope	6
9 Definition of audio presentation attributes	6
9.1 Shared presentation attributes	6
9.1.1 Clipping	6
9.2 Layout presentation attributes	7
9.2.1 Attenuation	7
9.2.2 Temporal envelope	7
9.2.3 Frequency envelope	8
9.2.4 Content architecture class	9
10 Definition of audio content portion attributes	9
10.1 Common coding attributes	9
10.1.1 Type of coding	9
10.2 Other coding attributes	10
10.2.1 Marker	10
10.3 Content information attribute	10

© ISO/IEC 1996

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from the publisher.

ISO/IEC Copyright Office • Case postale 56 • CH-1211 Genève 20 • Switzerland

Printed in Switzerland

11	Formal definitions of audio content architecture dependent data types	11
11.1	Introduction.....	11
11.2	Representation of document profile attributes	11
11.3	Representation of presentation attributes	11
11.4	Representation of coding attributes.....	12
11.5	Representation of non-basic features and non-standard defaults	13
12	Content layout process	13
12.1	Introduction.....	13
12.1.1	Purpose.....	13
12.1.2	Available area	13
12.1.3	Available time span.....	13
12.1.4	Presentation attributes	13
12.1.5	Audio content architecture classes	13
12.1.6	Association of content.....	14
12.2	Content layout process for audio content.....	14
13	Content presentation process.....	14
13.1	Introduction.....	14
13.2	Content presentation process for formatted processable form	14
14	Interactions with document architecture attributes.....	14
15	Definition of audio content architecture classes.....	16
15.1	Summary of audio presentation attributes.....	16
15.2	Summary of audio content portion attributes.....	16
Annex A	– SGML representation of audio content-specific attributes for ODL.....	17
A.1	Introduction.....	17
A.2	Names and public identifiers.....	17
A.3	Representation of attribute values.....	17
A.3.1	Constructed parameters.....	17
A.3.2	String parameters	17
A.3.3	Keyword parameters	17
A.3.4	Integer parameters.....	18
A.3.5	Real parameters.....	18
A.4	Presentation attributes	18
A.5	Coding attributes	19
Annex B	– Data type definitions for audio content architecture attributes	20
Annex C	– Summary of audio content architecture class	21
C.1	Formatted processable audio content architecture class	21
C.1.1	Presentation attributes	21
C.1.2	Content portion attributes.....	22
Annex D	– Summary of ASN.1 object identifiers.....	23
Annex E	– Modifications to other parts of ITU-T Rec. T.410-Series ISO/IEC 8613	24
E.1	Modifications to ITU-T Rec. T.411 ISO/IEC 8613-1	24
E.2	Modifications to ITU-T Rec. T.412 ISO/IEC 8613-2	24
E.3	Modifications to ITU-T Rec. T.414 ISO/IEC 8613-4	24
E.4	Modifications to ITU-T Rec. T.415 ISO/IEC 8613-5	24
Annex F	– Characteristics of audio encoding schemes	26

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. 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.

International Standard ISO/IEC 8613-9 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 18, *Document processing and related communication*, in collaboration with ITU-T. The identical text is published as ITU-T Recommendation T.419.

ISO/IEC 8613 consists of the following parts, under the general title *Information technology — Open Document Architecture (ODA) and interchange format*:

- *Part 1: Introduction and general principles*
- *Part 2: Document structures*
- *Part 3: Abstract interface for the manipulation of ODA documents*
- *Part 4: Document profile*
- *Part 5: Open Document Interchange Format*
- *Part 6: Character content architectures*
- *Part 7: Raster graphics content architectures*
- *Part 8: Geometric graphics content architectures*
- *Part 9: Audio content architectures*
- *Part 10: Formal specifications*
- *Part 11: Tabular structures and tabular layout*
- *Part 12: Identification of document fragments*
- *Part 13: Spreadsheet*
- *Part 14: Temporal relationship and non-linear structures*

Annexes A and E form an integral part of this part of ISO/IEC 8613. Annexes B to D and F are for information only.

Introduction

This ITU-T Recommendation | International Standard was prepared as a joint publication by TSS Study Group 8 and ISO/IEC Joint Technical Committee 1.

At present, the ITU-T Recommendations T.410-Series | International Standard ISO/IEC 8613 consists of:

- Introduction and general principles;
- Document structures;
- Abstract interface for the manipulation of ODA documents;
- Document profile;
- Open document interchange formats;
- Character content architectures;
- Raster graphics content architectures;
- Geometric graphics content architectures;
- Audio content architectures;
- Formal specification of the Open Document Architecture (FODA)
(the use of this Specification is applicable to ISO/IEC 8613 only);
- Tabular structures and tabular layout;
- Identification of document fragments;
- Temporal relationships and non-linear structures.

Further Recommendations | International Standards may be added to this series of ITU-T Recommendations | International Standard.

Development of this series of ITU-T Recommendations | International Standard was originally in parallel with ECMA-101: *Open Document Architecture*.

This ITU-T Recommendation | International Standard contains six annexes:

- Annex A (integral to ISO/IEC only): SGML representation of audio content-specific attributes for ODL;
- Annex B (non-integral): Data type definitions for audio content architecture attributes;
- Annex C (non-integral): Summary of audio content architecture class;
- Annex D (non-integral): Summary of ASN.1 object identifiers;
- Annex E (integral): Modifications to other common texts of ITU-T Rec. T.410-Series | ISO/IEC 8613;
- Annex F (non-integral): Characteristics of audio encoding schemes.

INTERNATIONAL STANDARD

ITU-T RECOMMENDATION

**INFORMATION TECHNOLOGY –
OPEN DOCUMENT ARCHITECTURE (ODA)
AND INTERCHANGE FORMAT:
AUDIO CONTENT ARCHITECTURES**

1 Scope

The purpose of this ITU-T Rec. T.410-Series | ISO/IEC 8613 is to facilitate the interchange of documents.

In the context of these Recommendations | International Standards, documents are considered to be items such as memoranda, letters, invoices, forms and reports, which may include pictures and tabular material. The content elements used within the documents may include graphic characters, geometric graphic elements, raster graphic elements and audio elements, all potentially within one document.

NOTE – These Recommendations | International Standards are designed to allow for extensions, such as hypermedia features, spreadsheets and additional types of content such as video.

In addition to the content types defined in these Recommendations | International Standards, ODA also provides for arbitrary content types to be included in documents.

These Recommendations | International Standards apply to the interchange of documents by means of data communications or the exchange of storage media.

These Recommendations | International Standards provide for the interchange of documents for either or both of the following purposes:

- to allow presentation as intended by the originator;
- to allow processing such as editing and reformatting.

The composition of a document in interchange can take several forms:

- formatted form, allowing presentation of the document;
- processable form, allowing processing of the document;
- formatted processable form, allowing both presentation and processing.

These Recommendations | International Standards also provide for the interchange of ODA information structures used for the processing of interchanged documents.

This Recommendation | International Standard:

- defines audio content architectures that can be used in conjunction with the document architecture defined in ITU-T Rec. T.412 | ISO/IEC 8613-2 and ITU-T Rec. T.424 | ISO/IEC 8613-14;
- defines those aspects of rendition applicable to the presentation of audio content;
- defines the presentation and content portion attributes applicable to these audio content architectures;
- describes an audio layout and presentation process, which, together with the document processing model described in ITU-T Rec. T.412 | ISO/IEC 8613-2, determines the layout of basic layout objects with associated audio content in the spatial and temporal dimensions of a document and the presentation of audio content on some suitable media.

2 Normative references

The following ITU-T Recommendations and International Standards contain provisions which through reference in this text, constitute provisions of this Recommendation | International Standard. At the time of publication, the editions indicated were valid. All Recommendations and Standards are subject to revision, and parties to agreements based on this Recommendation | International Standard are encouraged to investigate the possibility of applying the most recent edition of the Recommendations and Standards listed below. Members of IEC and ISO maintain registers of currently valid International Standards. The Telecommunication Standardization Bureau of the ITU-T maintains a list of currently valid ITU-T Recommendations.

2.1 Identical Recommendations | International Standards

- ITU-T Recommendation H.263¹⁾ | ISO/IEC 13818-3 ...¹⁾, *Information technology – Generic coding of moving pictures and associated audio information: Audio.*
- ITU-T Recommendation T.411 (1993) | ISO/IEC 8613-1:1994, *Information technology – Open Document Architecture (ODA) and interchange format: Introduction and general principles.*
- ITU-T Recommendation T.412 (1993) | ISO/IEC 8613-2:1995, *Information technology – Open Document Architecture (ODA) and interchange format: Document structures.*
- ITU-T Recommendation T.414 (1993) | ISO/IEC 8613-4:1994, *Information technology – Open Document Architecture (ODA) and interchange format: Document profile.*
- ITU-T Recommendation T.415 (1993) | ISO/IEC 8613-5:1994, *Information technology – Open Document Architecture (ODA) and interchange format: Open Document Interchange Format.*
- ITU-T Recommendation T.416 (1993) | ISO/IEC 8613-6:1994, *Information technology – Open Document Architecture (ODA) and interchange format: Character content architectures.*
- ITU-T Recommendation T.417 (1993) | ISO/IEC 8613-7:1994, *Information technology – Open Document Architecture (ODA) and interchange format: Raster graphics content architectures.*
- ITU-T Recommendation T.418 (1993) | ISO/IEC 8613-8:1994, *Information technology – Open Document Architecture (ODA) and interchange format: Geometric graphics content architectures.*
- ITU-T Recommendation T.421 (1994) | ISO/IEC 8613-11:1995, *Information technology – Open Document Architecture (ODA) and interchange format: Tabular structures and tabular layout.*
- ITU-T Recommendation T.422 (1995) | ISO/IEC 8613-12:1996, *Information technology – Open Document Architecture (ODA) and interchange format: Identification of document fragments.*
- ITU-T Recommendation T.424²⁾ | ISO/IEC 8613-14 ...²⁾, *Information technology – Open Document Architecture (ODA) and interchange format: Temporal relationships and non-linear structures.*

2.2 Paired Recommendations | International Standards equivalent in technical content

- CCITT Recommendation X.208 (1988), *Specification of Abstract Syntax Notation One (ASN.1).*
ISO/IEC 8824:1990, *Information technology – Open Systems Interconnection – Specification of Abstract Syntax Notation One (ASN.1).*

2.3 Additional references

- ISO 8879:1986, *Information processing – Text and office systems – Standard Generalized Markup Language (SGML).*
- ISO 266:1975³⁾, *Acoustics – Preferred frequencies for measurements.*
- CCITT Recommendation G.711 (1988), *Pulse Code Modulation (PCM) of voice frequencies.*
- CCITT Recommendation G.721 (1988), *32 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM).*
- CCITT Recommendation G.722 (1988), *7 kHz audio-coding within 64 kbit/s.*

¹⁾ Presently at the stage of draft.

²⁾ To be published.

³⁾ Currently under revision.

- CCITT Recommendation G.723 (1988), *Extensions of Recommendation G.721 adaptive differential pulse code modulation to 24 and 40 kbit/s for digital circuit multiplication equipment application.*
- CCITT Recommendation G.726 (1990), *40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM).*
- CCITT Recommendation G.727 (1990), *5-, 4-, 3-, and 2-bits/sample embedded Adaptive Differential Pulse Code Modulation (ADPCM).*
- CCITT Recommendation G.728 (1992), *Coding of speech at 16 kbit/s using low-delay code excited linear prediction.*
- CCITT Recommendation J.41 (1988), *Characteristics of equipment for the coding of analogue high quality sound-program signals for transmission on 384 kbit/s channels.*
- CCITT Recommendation J.42 (1988), *Characteristics of equipment for the coding of analogue medium quality sound-program signals for the transmission on 384 kbit/s channels.*
- ISO/IEC 11172-3:1993, *Information technology – Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s – Part 3: Audio.*
- IEC 899:1987, *Sampling rate and source encoding for professional digital audio recording.*
- I-ETS 300 036 *European digital cellular communications system (phase 1); full rate speech transcoding (GSM 06-10).*
- ANSI S4.28-1984, *Preferred sampling frequencies for professional digital audio applications employing pulse code modulation.*
- ANSI S4.40-1985, *Serial transmission format for linearly represented digital audio data.*
- ANSI/AES 3-1992, *Serial transmission format for 2-channel linearly represented audio data.*

3 Definitions

For the purposes of this Recommendation | International Standard, the definitions given in ITU-T Rec. T.411 | ISO/IEC 8613-1 apply.

For the purpose of this Recommendation | International Standard, the following definitions also apply.

3.1 sampled sound: A sequence of digital values that represent instantaneous pressure of an auditory sensation taken at discrete intervals of time.

3.2 required duration: The time required to present the audio content information derived by the audio content layout process taking into account applicable presentation and coding attributes.

4 Abbreviations

For the purposes of this Recommendation | International Standard, the abbreviations given in ITU-T Rec. T.411 | ISO/IEC 8613-1 apply.

5 Conventions

For the purposes of this Recommendation | International Standard, the conventions given in ITU-T Rec. T.411 | ISO/IEC 8613-1 apply.

6 General principles

6.1 Content architectures

This Recommendation | International Standard defines one class of audio content architecture:

- A formatted processable form, which allows for document content to be processed and also to be presented as intended by the originator; formatted processable content can be associated with any basic component.

6.1.1 Formatted processable content architecture class

Formatted processable audio content is intended to be laid out in the temporal dimension, reformatted in the temporal dimension or presented by the recipient in accordance with the originator's intent. This form of content may be used in formatted, processable and formatted processable form documents.

6.2 Content

The content of a basic component that is structured according to an audio content architecture is a sequence of sound samples. These sound samples are represented by a sequence of octets in accordance with a particular coding scheme for audio information. A given encoding scheme can be characterized by features such as the sampling rate, number of bits per sample, the number of channels sampled and compression technique used.

If a content portion containing audio content is associated with a basic component, no other content portion shall be associated with that basic component.

6.3 Presentation attributes

Presentation attributes are applicable to basic components and specify information for presenting the content of the basic component and are defined in clause 9.

6.4 Content portion attributes

Content portion attributes are applicable to content portions and specify information related to the identification and coding of the content. Content portion attributes are defined in clause 10.

6.5 Coding of content information

The content information is encoded according to one of the types of coding indicated in the content portion coding attribute "type of coding". Other coding attributes are defined in 10.2.

7 Audio positioning

A basic logical component with a formatted processable form content architecture class must undergo the content layout process before its associated content can be positioned and presented.

During the document layout process, audio content associated with basic logical components is positioned within lowest level frames in the specific layout structure.

For a basic layout object with associated audio content, the value of the attribute "dimensions" is 'null'. The document layout process inserts a basic layout object with associated audio content at the same position in the specific layout structure as a layout object with visual content. Basic layout objects with associated audio content have no horizontal or vertical extent and the layout directives "block alignment", "concatenation", "fill order", "offset", "separation" and "synchronization" have default values, as specified in ITU-T Rec. T.412 | ISO/IEC 8613-2, when applied to such basic layout objects.

The content layout process (defined in clause 12) determines the required duration of the presentation of the audio content.

7.1 Basic concepts

7.1.1 Measurement units

For audio content, the unit for positioning elements of sampled sound on the time axis is the Scaled Time Unit (STU) as specified in E.3.1 of ITU-T Rec. T.424 | ISO/IEC 8613-14.

The sampling rate of audio content information is specified in one of the following ways:

- 1) the coding method specified by the attribute "type of coding" uniquely determines the sampling rate; or
- 2) the content information itself determines the sampling rate in the way specified by the coding method.

7.2 Positioning of elements of sampled sound in a basic layout time frame

The exact position in time, relative to the start of the presentation of the basic layout object, of the n th element of the sampled sound is determined by the following formula:

$$(n/F) - s \text{ STUs after the start of the presentation of the basic layout object}$$

provided that:

$$n/F < d, (n/F - s) \geq 0 \text{ and } n/F < e$$

where

F is the sampling rate – number of samples per STU;

s is the clipping start time in STUs;

e is the clipping end time in STUs; and

d is the duration of the original sample in STUs.

7.3 Reference recording level

The “reference recording level” is a document profile attribute that allows a recipient to determine the actual differences between changes in the presentation levels specified by an originator rather than the ratio of changes in presentation levels. The specification of the document profile attribute “reference recording level” is specified in Annex E.

8 Audio presentation

Two groups of specifications may apply to the presentation of basic layout objects with audio content, namely those relating to:

- attenuation and temporal envelope;
- frequency envelope.

These groups of specifications are defined below. In addition, it is possible to specify a sequence of marker labels by means of the attribute “marker”. This allows specific portions of the audio information to be identified by associating a label with a selected portion specified by a time offset.

The specification of attenuation, temporal and frequency envelope can be associated with specific channels. A *channel* represents the means of presenting audio information and is identified by its channel name.

NOTE – Typical channel names might be “left” and “right” to identify the two aspects of stereophonic presentation.

When the content information of an audio content portion is presented, a clipping specification can identify which part of that content information is intended to be presented. The specification for clipping can identify a portion relative to the start or the end of the content information.

The specifications of temporal envelope and clipping specify either time offsets or marker labels (object identifiers or character strings) which reference markers containing time offsets. Such labels can be more meaningful to end-users rather than an integer specifying a particular time offset.

8.1 Attenuation and temporal envelope

Changes in the presentation level of all of the recorded audio information or parts of it may be specified with respect to a reference recording level value (see 7.3). Changes in the presentation level of the whole of the recorded audio information are controlled by the presentation attribute “attenuation”.

NOTE 1 – Changes to the presentation level of audio information can be used to achieve end-user features such as *volume*.

Changes to the presentation level of parts of the audio information are controlled by the presentation attribute “temporal envelope” where for each time interval in a sequence of intervals, an associated change in the presentation level can be specified for a set of channels. The sequence of time intervals is specified relative to the start of the audio information.

NOTE 2 – Such changes to the presentation level can be used to achieve end-user features such as *fade-in* and *fade-out*.