

STANDARDS AND INFORMATION DOCUMENTS

AES41-1-2012
rev. AES41-2009
reaffirmed 2017



AES standard for digital audio - Audio-embedded metadata - Part 1: General

Users of this standard are encouraged to determine if they are using the latest printing incorporating all current amendments and editorial corrections. Information on the latest status, edition, and printing of a standard can be found at:
<http://www.aes.org/standards>

AUDIO ENGINEERING SOCIETY, INC.
551 Fifth Avenue, New York, New York 10176, US.



The AES Standards Committee is the organization responsible for the standards program of the Audio Engineering Society. It publishes technical standards, information documents and technical reports. Working groups and task groups with a fully international membership are engaged in writing standards covering fields that include topics of specific relevance to professional audio. Membership of any AES standards working group is open to all individuals who are materially and directly affected by the documents that may be issued under the scope of that working group.

Complete information, including working group scopes and project status is available at <http://www.aes.org/standards>. Enquiries may be addressed to standards@aes.org

The AES Standards Committee is supported in part by those listed below who, as Standards Sustainers, make significant financial contribution to its operation.



This list is current as of 2018/4/18

AES standard for digital audio - Audio-embedded metadata - Part 1: General

Published by

Audio Engineering Society, Inc.

Copyright © 2009, 2012, 2017 by the Audio Engineering Society

Abstract

AES41 provides for the carriage of audio metadata by embedding it in the audio samples themselves. This tightly associates the metadata with the audio, yet makes it fragile so that changes to the audio will invalidate the metadata. Several metadata sets have been defined, covering applications such as cascaded compression (bit-rate reduction), and loudness control.

This Part specifies the method by which a metadata set is embedded in the audio samples. Other parts define the content and format of the metadata sets.

An AES standard implies a consensus of those directly and materially affected by its scope and provisions and is intended as a guide to aid the manufacturer, the consumer, and the general public. The existence of an AES standard does not in any respect preclude anyone, whether or not he or she has approved the document, from manufacturing, marketing, purchasing, or using products, processes, or procedures not in agreement with the standard. Prior to approval, all parties were provided opportunities to comment or object to any provision. Attention is drawn to the possibility that some of the elements of this AES standard may be the subject of patent rights other than those identified herein. AES shall not be held responsible for identifying any or all such patent rights. Approval does not assume any liability to any patent owner, nor does it assume any obligation whatever to parties adopting the standards document. Recipients of this document are invited to submit, with their comments, notification of any relevant patent rights of which they are aware and to provide supporting documentation. This document is subject to periodic review and users are cautioned to obtain the latest edition.

2017-12-31 printing

Contents

0 Introduction	5
0.1 Rationale for this standard	5
0.2 Conventions used in this standard	5
0.3 Patents	5
1 Scope	6
2 Normative references	6
3 Definitions, symbols, and abbreviations	6
3.1 Definitions	6
3.2 Symbols	6
3.3 Abbreviations	7
4 Method for describing the bit-stream syntax	8
5 Specification of the audio coder control data bit-stream syntax	9
6 Semantics of the bit-stream syntax	10
6.1 Header	10
6.2 AES41 metadata type registry	10
6.3 Error_check	11
6.4 Ancillary_data	11
7 Method of signaling the audio coder control data bit stream on the AES3 interface	11
Annex A (informative) AES41 metadata types	13
Annex B Bibliography	14

Foreword

[This foreword is not a part of *AES standard for digital audio — Recoding data set for audio bit-rate reduction*, AES41-2000.]

This document describes a detailed specification of a set of coding-decision data for MPEG Layer II with a generic transport mechanism developed as project AES-X78 in the AESSC SC-02-02 Working Group on Digital Input-Output Interfacing of the SC-02 Subcommittee on Digital Audio. The task group was lead by A. Mason. The intent has been to produce a standard in cooperation with other bodies. Individual experts associated with potential end-users, equipment developers, and manufacturers who are directly and materially affected by this project were encouraged to participate in the process of standardization by joining SC-02-02.

In the near future, digital compression techniques are expected to dominate the broadcast television environment. Digital compression will be used for a significant amount of program acquisition, for program production and storage, and for distribution and broadcasting.

Compression systems add impairments to sound. The audibility of these impairments depends on the degree of compression and the techniques used. It has been proved that re-encoding with minimal further impairment can be achieved by using information describing the previous encoding process. The information that is preserved varies according to the encoding scheme that has been used. For example, for transparent cascaded coding of MPEG-1 Layer II or MPEG-2 Layer II audio, the information required by the downstream encoder includes:

- . the positions of the frame boundaries;
- . the bit rate of the compressed bit stream before decoding;
- . the coding mode (monophonic, joint stereo, and so on);
- . the bit allocation for each sub-band within a frame;
- . the scale factors for each sub-band within each sub-block of a frame.

In order to help with the process of editing decoded audio, it is useful to send additional coding-decision data to describe any timing offset that may be introduced by editing only at points corresponding to block boundaries (for example, 24-ms boundaries for MPEG-1 Layer II at 48 kHz). These data are particularly relevant for editing sound associated with co-timed video. The bit rate of these additional data is less than 30 kbit/s. A cyclic-redundancy check word must be appended to frames of data to provide a means to detect errors in the data.

This data signal can be used by a downstream encoder to guide its coding decisions. For example, if the audio signal is changed in any way (as it would by mixing it with another signal) then it could be inappropriate for the downstream encoder to re-use any previous coding decisions. However, in such a case, when a suitable transport mechanism is used, the coding-decision data would be corrupted so the coder could detect that the data are no longer valid.

This specification describes in detail the content of the coding-decision information, such as MPEG-header information, bit allocations, and scale factors, plus its representation in terms of the number, order, and meaning of its bits. A mechanism by which these data can be conveyed in a pulse-code-modulation audio signal is also described. The mechanism modifies one of the least significant bits of the audio word to signal the data. The precise time relationship of the data to the audio is also defined.

R. Finger, chair, SC-02-02
J. Dunn, vice-chair, SC-02-02
2000-04-08