



## Administrative Metadata for Digital Audio Files

This data dictionary provides element-by-element descriptions of the <audioMetadata> block specified in the Harvard University Library Digital Repository Services (DRS) XML DTD Specification. As such, it represents **part** of the comprehensive list of technical data elements relevant to the management of digital audio. (See “Usage Note” below.) In this context, “management” refers to the tasks and operations needed to support audio quality assessment and audio data processing throughout the audio life cycle. “Quality assessment” is defined broadly, as it refers both to machine operations and curatorial evaluations. Technical metadata for digital audio files have been identified to anchor meaningful attributes of audio quality that can be measured objectively, such as bit depth and sample rate. In addition, data elements have been proposed to support subjective assessments of “current value” by digital repository managers, curators, or audio specialists seeking to determine whether intrinsic quality (aesthetic or functional) sufficiently justifies associated maintenance or processing costs.

Data elements and values are drawn from industry specifications for file formats, as well as published guidelines from other institutions seeking to store and manage large audio collections. (See, *Sources Consulted*.) Elements and values in this document apply inclusively to all digital audio formats, regardless of whether they are used for archival or delivery functions.

Usage Note: Consult [DRS User Manual for Data Loading](http://hul.harvard.edu/ois/systems/drs/drs_load_manual.pdf) <[http://hul.harvard.edu/ois/systems/drs/drs\\_load\\_manual.pdf](http://hul.harvard.edu/ois/systems/drs/drs_load_manual.pdf)> for additional instructions needed to construct a valid XML batch file for DRS deposit. This corresponding documentation provides deposit instructions, the full DTD specification, and element-by-element descriptions for XML blocks other than <audioMetadata>. For example, characteristics such as <mimetype> and <signature> are applicable across formats, and so occur in base metadata rather than within <audioMetadata>.

Note that for audio files in the DRS additional metadata and auxiliary files in the form of edit decision lists, waveform reduction files, platform-specific processing files, and processing history files, may also be submitted in order to assist in preservation and the production of new delivery files. Most of these will be encapsulated within a METS file describing the audio. For details, contact the LDI Metadata Analyst listed below.

September 4, 2007

© President and Fellows of Harvard College

<http://hul.harvard.edu/ois/systems/drs/audiometadata.pdf>

<b>Document Administration.....</b>	<b>3</b>
<b>Digital Audio: Technical Metadata.....</b>	<b>3</b>
Section 1: Audio Attributes.....	3
Duration.....	3
Bit depth.....	3
Sample rate.....	3
Number of channels.....	4
Audio data format type.....	4
Audio data orientation.....	4
Sound channel map.....	5
Sound field.....	5
Sound channel assignment.....	5
Channel number.....	5
Sound map location.....	5
Bit Rate Reduction.....	6
Codec Name.....	6
Codec Name Version.....	6
Codec Creator Application.....	7
Codec Creator Application Version.....	7
Codec Quality.....	7
Data Rate.....	7
Data Rate Mode.....	8
Offset to first relevant sample.....	8
Blocksize.....	8
First valid byte in block.....	8
Last valid byte in block.....	8
Word size.....	9
Time Stamp Start.....	9
Time Stamp End.....	9
<b>Sources Consulted.....</b>	<b>9</b>
<b>Revision History.....</b>	<b>11</b>

## Document Administration

This document is authored and administered by the LDI Technical Team. Values and mappings will be expanded as local needs for additional file formats develop. All questions and comments should be directed to the LDI Metadata Analyst (Robin Wendler, <[r\\_wendler@harvard.edu](mailto:r_wendler@harvard.edu)> 617-495-3724).

## Digital Audio: Technical Metadata

### Section 1: Audio Attributes

Values for these data elements are constructed to facilitate parsing and reporting. Data in the mandatory (#1-7.2b) and mandatory-if-applicable (#8-8.7) fields will be used to process files in order to produce new deliverables or to migrate archival files. Optional data (fields #9-15) are to be collected to support audio processing, in some cases, *and* to generate summary “collection assessment reports” for curators and data managers.

Number	1
Name	Duration
Definition	<i>The length of sound in Time-code Character Format (TCF).</i>
Required	M
Repeatable	N
Values	Any valid Time-code Character Format value.
Mapping	
Examples	00:10?25.14
Notes	TCF may require the use of indicator codes '<' and '&', which cannot be entered directly in XML format. If these codes are required they must be entered into an XML document using the <code>&amp;lt;</code> and <code>&amp;amp;</code> entities.  See AES31-3-1999, <i>AES standard for network and file transfer of audio – Audio-file transfer and exchange – Part 3: Simple project interchange</i>

Number	2
Name	Bit depth
Definition	<i>Designates the number of bits per sample point of the audio data.</i>
Required	M
Repeatable	N
Values	Integer values, generally 8, 16, 24, or 32
Mapping	
Examples	
Notes	A sample point represents one sample of one channel

Number	3
Name	Sample rate
Definition	<i>Designates the number of samples per second in sample frames.</i>
Required	M

Repeatable	N
Values	Typical values are 32000 , 44100, 48000, 88200, 96000, 192000
Mapping	
Examples	
Notes	

Number	4
Name	Number of channels
Definition	<i>Designates the number of audio channels present.</i>
Required	M
Repeatable	N
Values	Whole numbers
Mapping	
Examples	1 2 3 4
Notes	

Number	5
Name	Audio data format type
Definition	The format used in the digitizing process
Required	M
Repeatable	N
Values	PCM DSD MPEG1 MPEG2
Mapping	
Examples	
Notes	

Number	6
Name	Audio data orientation
Definition	Designates the byte order of the data within the audio data chunk (that is, not of the entire file).
Required	M
Repeatable	N
Values	0 = Big endian (most significant byte first) 1 = Little endian (least significant byte first)
Mapping	
Examples	
Notes	

Number	7
Name	Sound channel map
Definition	Provides a map of one or more sound channel assignments
Required	M
Repeatable	R
Values	Contains a Sound field element and one or more Sound channel assignment elements
Mapping	
Examples	
Notes	

Number	7.1
Name	Sound field
Definition	<i>Aural space on source recording.</i>
Required	M
Repeatable	N
Values	mono stereo surround
Mapping	
Examples	
Notes	

Number	7.2
Name	Sound channel assignment
Definition	<i>Maps an audio channel/track to its intended aural position/destination.</i>
Required	M
Repeatable	R
Values	Two mandatory subelements: <b>Channel number</b> and <b>Sound map location</b>
Mapping	
Examples	
Notes	

Number	7.2.a
Name	Channel number
Definition	<i>Identifies the number of an audio channel.</i>
Required	M
Repeatable	N
Values	Integers
Mapping	
Examples	0 1
Notes	

Number	7.2.b
Name	Sound map location

Definition	<i>Maps a sound channel to its destination</i>
Required	M
Repeatable	N
Values	left right center left_rear right_rear sub timecode
Mapping	
Examples	
Notes	

Number	8
Name	Bit Rate Reduction
Definition	<i>Codec used to produce a bit rate reduced version of the audio file.</i>
Required	MA
Repeatable	R
Values	Three subelements: Name, Creator, Quality, which are all required, non-repeatable within the context of one Bit Rate Reduced Codec instance
Mapping	
Examples	
Notes	

Number	8.1
Name	Codec Name
Definition	<i>Codec used to produce a bit rate reduced version of the audio file.</i>
Required	MA
Repeatable	N
Values	codec name
Mapping	
Examples	16 kbps Voice 24 kbps Stereo Music
Notes	

Number	8.2
Name	Codec Name Version
Definition	<i>The version of the named Codec used to produce a bit rate reduced version of the audio file.</i>
Required	MA
Repeatable	N
Values	codec name version
Mapping	
Examples	

Notes	
-------	--

Number	8.3
Name	Codec Creator Application
Definition	<i>Application used to encode audio</i>
Required	MA
Repeatable	N
Values	application name
Mapping	
Examples	Real Audio Producer
Notes	

Number	8.4
Name	Codec Creator Application Version
Definition	<i>Version of the application used to encode audio</i>
Required	MA
Repeatable	N
Values	application name version
Mapping	
Examples	Version 5 Macintosh
Notes	

Number	8.5
Name	Codec Quality
Definition	<i>Indicates if the codec in use is 'code regenerating' or lossy</i>
Required	MA
Repeatable	N
Values	0 = code regenerating 1 = lossy
Mapping	
Examples	
Notes	

Number	8.6
Name	Data Rate
Definition	<i>Average bytes per second</i>
Required	MA
Repeatable	N
Values	
Mapping	
Examples	
Notes	Relevant specifically for non-fixed-rate audio. If present, Data Rate Mode is required.

Number	8.7
Name	Data Rate Mode
Definition	<i>Designates whether the data rate is fixed or variable</i>
Required	MA
Repeatable	N
Values	0 = Fixed 1 = Variable
Mapping	
Examples	
Notes	

Number	9
Name	Offset to first relevant sample
Definition	<i>Designates the offset in bytes to the first sample frame in the audio data.</i>
Required	O
Repeatable	N
Values	Any positive integer
Mapping	
Examples	
Notes	

Number	10
Name	Blocksize
Definition	<i>Designates the size in bytes of a block to which the audio is aligned.</i>
Required	O
Repeatable	N
Values	Any positive integer or zero
Mapping	
Examples	
Notes	

Number	11
Name	First valid byte in block
Definition	<i>Designates the first byte in each block in which valid sound data occurs.</i>
Required	O
Repeatable	N
Values	Any positive integer between zero and the value of <b>BlockSize</b>
Mapping	
Examples	
Notes	

Number	12
Name	Last valid byte in block
Definition	<i>Designates the last byte in each block in which valid sound data occurs.</i>

Required	O
Repeatable	N
Values	Any positive integer between <b>First valid byte in block</b> and the value of <b>BlockSize</b>
Mapping	
Examples	
Notes	

Number	13
Name	Word size
Definition	<i>Designates the number of bytes per audio word.</i>
Required	O
Repeatable	N
Values	Integer values, generally 1, 2, 3, or 4
Mapping	
Examples	
Notes	Used when the word size differs from the bit depth

Number	14
Name	Time Stamp Start
Definition	<i>The designated time value of the first sample of the file expressed in Time-code Format (TCF).</i>
Required	O
Repeatable	N
Values	Any valid TCF value
Mapping	
Examples	01:20:30.10
Notes	See Note 1

Number	15
Name	Time Stamp End
Definition	<i>The designated time value of the last sample of the file expressed in Time-code Format (TCF).</i>
Required	O
Repeatable	N
Values	Any valid TCF value
Mapping	
Examples	01:20:30.10
Notes	

## Sources Consulted

AES31-3-1999: AES standard for network and file transfer of audio –

Audio-file transfer and exchange – Part 3: Simple project interchange (PDF, 228KB)  
<http://aessec.aessc.aes.org/pub/aes31-3-1999.pdf>

AES Working Draft: Administrative Metadata for Audio Objects Part I: Audio Processing History Metadata Schema

AES Working Draft: Audio Object Metadata

Apple Computer, Inc., Audio Interchange File Format: A Standard for Sampled Sound Files, Version 1.3, January 4, 1989 (PDF, 58KB)

<<http://preserve.harvard.edu/standards/audioiffspecification1-3.pdf>>

Sonic System Audio File Format. Roy Waldspurger, Sonic Solutions

CONFIDENTIAL

## Revision History

VERSION	CONTACT	LAST UPDATE	STATUS
1.0	RW	April 11,2002	released for use
1.1	RW	August 1, 2003	Removed <Number of sample frames> element Made <Sound channel map> Repeatable Made <Sound field> a sub-element of <sound channel map> Made <Data rate> and <Data rate mode> sub-elements of <Bit rate reduction> Renamed <Bit rate reduced codec> to <Bit rate reduction> and expanded number of sub-elements from three <Codec name, Codec creator, and Codec quality> to seven (see 8.1-8.7)
1.2	SA/RW	February 11, 2004	Added note for TCF Clarified <Sound channel map> value statement