Federal Agencies
Digital Guidelines Initiative

Guidelines for Embedding Metadata in Broadcast WAVE Files

Version 3
April 26, 2021

The FADGI Audio-Visual Working Group
http://www.digitizationguidelines.gov/audio-visual/

Guidelines for Embedding Metadata in Broadcast WAVE Files

By the Federal Agencies Digital Guidelines Initiative (FADGI)
Audio-Visual Working Group
http://www.digitizationguidelines.gov/audio-visual/

Version 3

Approved by Working Group: April 26, 2021

What is this document?

This is a revision of the Federal Agencies Digital Guidelines Initiative (FADGI) Audio-Visual Working Group Broadcast WAVE file metadata recommendation first published in September 2009 with a second version published in 2012.See below for the change list for this third version published in 2021. It is one of four documents pertaining to the embedding of metadata in digital audio files. The companion documents are:

- *Introductory Discussion for the Proposed Federal Agencies Guideline (updated April 2012).* http://www.digitizationguidelines.gov/audio-visual/documents/Embed_Intro_20120423.pdf
- *Consultant's report on embedding options in digital audio files (June 2009).* http://www.digitizationguidelines.gov/audio-visual/documents/AVPS_Audio_Metadata_Overview_090612.pdf
- *Discussion paper: Identifiers: Types and Characteristics (November 2011).* http://www.digitizationguidelines.gov/audio-visual/documents/IdentifiersTypesCharacteristics_20111121.pdf

The first part of this document recommends actions pertaining to data elements in the BEXT chunk established as part of the Broadcast WAVE (BWF) file specification by the European Broadcast Union, and references are made to this specification's three versions:

- BWF Version 0. The specification of the Broadcast WAVE Format for PCM audio data (now referred to as Version 0) was published in 1997 as EBU Tech 3285.
- BWF Version 1. Version 1 differs from Version 0 only in that 64 of the 254 reserved bytes in Version 0 are used to contain a SMPTE UMID. Published July 2001.
- BWF Version 2. Version 2 is a substantial revision of Version 1 which incorporates loudness metadata (in accordance with EBU R 128) and which takes account of the publication of Supplements 1 – 6 and other relevant documentation. Published May 2011.

Change list from version 2, published in 2012 (superseded by this document)

- Simplified title from “Embedding Metadata in Digital Audio Files: Guideline for Federal Agency Use of Broadcast WAVE Files” to “Guidelines for Embedding Metadata in Broadcast WAVE Files”
- For OriginationDate and ICRD, FADGI adds the option for a date range using double hyphens [--] as the interval designator, such as 2002--2003, in place of a solidus (a.k.a. forward slash "/") although a solidus is also permitted.
- Changes in CodingHistory:
    - Adding values for sampling frequency, bit depth and mode beyond what is listed in EBU R98 to reflect current needs and tool capabilities
        - Added values for sampling frequency: 96000, 176400, 192000, 384000, 768000 Hz
        - Added value for word length/bit depth: 32 bits.
        - Added values for mode: multitrack, multichannel and other.
    - Mode definition: To clarify the use of this field in CodingHistory, FADGI states that ‘mode’ is a synonym for ‘sound field’ and includes definitions for allowable mode values.
    - Definitions for all values in mode: mono, stereo, dual mono, joint stereo, multitrack, multichannel and other
    - Updated examples for coding history to reflect other changes
- Addition of guidance for use of Associated Data List Chunk (<adtl>), Label Chunk (<labl>),Note Chunk (<note>) and Labeled Text Chunk (<ltxt>) to accompany use of Cue Chunk (<cue>) and Cue Points.
- Correction from “LIST info chunk” to “INFO List Chunk” to comply with published specification
- Updated Specification and Reference list in Part V

Table of Contents

PART I. THE BEXT CHUNK

I.A Strongly recommended elements

If the Working Group had the authority to do so, these would be "required."

Originator	
EBU Tech 3285	ASCII string (maximum 32 characters) containing the name of the originator/producer of the audio file. If the length of the string is less than 32 characters, the field is ended by a null character. (Established in version 0 of the BWF specification.)
FADGI Application	This element contains the entity responsible for the creation, maintenance, preservation of this digital item. Entity designations should be as specific as possible including a two-character country code to avoid the potential for conflict in the responsible organization's name. If space permits within the 32 character limit, the archival entity should be identified at the most specific level within the institution. Use a standard abbreviation of entity names such as those found in the Guide to Government Acronyms & Abbreviations. If an entity is not on this list, use a familiar abbreviation. Use the standard two-character ISO 3166 alpha 2 country code list.
Char limit	32
Mandatory/optional	Strongly recommended (if the Working Group had authority: "required")
Values	[Country code]comma space[Entity name]
Example LC	US, LOC/RSS [RSS = Recorded Sound Section]
Example NARA	US, NARA
Example EPA	US, EPA

OriginatorReference	
EBU Tech 3285	ASCII string (maximum 32 characters) containing a non ambiguous reference allocated by the originating organization. If the length of the string is less than 32 characters, the field is ended by a null character. (Established in version 0.)

FADGI Application	This element contains the principal identifier or the "best" identifier which uniquely differentiates one object from another, preferably at the file level. If the principal identifier string is less than 32 characters, enter the entire identifier string. If the principal identifier string is longer than 32 characters, enter this text: "See Description for identifiers." Do not embed identifiers that could pose a possible security risk, e.g., by exposing exact pathnames. NOTE: The Working Group perceived value in the practice of repeating the principal identifier as the first identifier in the BEXT Description element (which has 256 available characters) but did not make this detail a strong recommendation. Comments from readers are welcome.
Char limit	32
Mandatory/optional	Strongly recommended (if the Working Group had authority: "required")
Values	Under 32 characters: Identifier string Over 32 characters: See Description
Example NARA	See Description for identifiers [Explanation: actual identifiers extend more than 32 characters.]
Example LC	1201566-2-1 [Explanation: unique number generated by the MAVIS collections management database. Preferred identifier for the LC Recorded Sound Section. Generally for local use.]

Description	
EBU Tech 3285	ASCII string (maximum 256 characters) containing a free description of the sequence. To help applications which only display a short description, it is recommended that a résumé of the description is contained in the first 64 characters, and the last 192 characters are use for details. If the length of the string is less than 256 characters, the last one is followed by a null character (00). (Established in version 0.)

FADGI Application	This element is recommended as a container for identifiers for the work at hand and/or as pointers to additional, non-embedded (externally maintained) metadata. Members of the Working Group have repeatedly encountered the need to provide multiple identifiers for a given item. The resulting extent of data cannot be accommodated in the OriginatorReference element. For these reasons, the Working Group's recommendations for the Description element deviate from the EBU specification. In some cases, the 256-character limit will prevent an agency from listing all of its identifiers; the most important or helpful should be provided. Do not embed identifiers that could pose a possible security risk, e.g., by exposing exact pathnames. NOTE: The Working Group perceived value in two practices but wished to leave these as optional. The first is the tagging of identifiers (see examples, typically URLs) to permit them to be properly understood. The second is the repetition of the principal identifier (as provided without tagging) in OriginatorReference as the first identifier in Description, where labeling as to its origin or purpose can be provided. Comments welcome.
Char limit	256
Mandatory/optional	Strongly recommended (if the Working Group had authority: "required")
Values	If labeled: Identifier [comma space] type [comma space] comment [semicolon-space if more than one identifier] If no labeling: Identifier
Example NARA	[Two labeled identifiers] 58979818, local, principal ID original filename; 306-MUSA-9658B, local, RG-Series-Item Number
Example LC	[One labeled identifier] http://hdl.loc.gov/loc.mbrsmi/westhpp.2033, URL, principal ID handle
Example LC	[One unlabeled identifier] http://hdl.loc.gov/loc.mbrsmi/westhpp.2033
Example LC	[One unlabeled identifier] 1201566-2-1 [Explanation: unique number generated by the MAVIS collections management database. Generally for local use.]
Example LC	[One unlabeled identifier] RYI_6039 [Explanation: Recorded Sound Section shelf number for the original physical item that has been digitally reformatted. For local use.]
Example LC	[One unlabeled identifier] Harmonia Mundi France HM 957 [Explanation: Label information for a phonodisc in the Recorded Sound Section collection]
Example California Digital Library	[One labeled identifier] ark.cdlib.org.org/ark:/13030/tf5p30086k, URL, ARK

Example LC	[One labeled identifier] http://lccn.loc.gov/mp76000002, URL, Permalink

OriginationDate	
EBU Tech 3285	Ten ASCII characters containing the date of creation of the audio sequence. (Established in version 0.)
FADGI Application	This element contains the file creation date. This is understood to mean the local date in the timezone for the archival entity; the structure of the bext chunk does not permit ISO 8601 datetime indication, which unambiguously indicates date and time in terms of UTC (Coordinated Universal Time or Temps Universel Coordonné).
Char limit	10
Mandatory/optional	Strongly recommended (if the Working Group had authority: "required")
Values	ISO 8601 YYYY-MM-DD. Year is defined from 0000 to 9999; Month is defined from 01 to 12 (use leading zeroes if less than 10); Day is defined from 01 to 28, 29, 30 or 31 (use leading zeroes if less than 10). The separator between the items is a hyphen [-] with no spaces. Thus the string length may be four, seven, or ten. If the string entered is less than ten characters, end it with a null character **Note:** the Working Group adheres to the ISO 8601 truncation convention: the year should always be given but, if unknown, values for month and day are omitted. When the date is represented by a range, such as 2008--2009, double hyphens [--] may be used as the interval designator in place of a solidus (a.k.a. forward slash "/").
Example	2005-11-09

Version	
EBU Tech 3285	An unsigned binary number giving the version of the BWF. The number is particularly relevant for the carriage of the UMID and loudness information. This element contains the BWF version. For Version 1 it shall be set to 0001h and for Version 2 it shall be set to 0002h. Version 2 is backwards compatible with Versions 1 and 0. This means that software designed to read Version 1 and Version 0 files will interpret the files correctly except that Version 0 software will ignore the UMID and loudness information which may be present and Version 1 software will ignore the loudness information. Therefore, users of such devices will lose metadata unless special precautions are taken. In addition, early BWF- aware devices will be unable to cope with the larger RF64 and MBWF files and may not recognise any of the chunks which have been defined since 2001. The change is also forwards compatible. This means that Version 2 software will be able to read Version 0 and Version 1 files correctly. Software needs to read the <Version> field to determine if a UMID and loudness metadata are present. (Established in version 0, extended in later versions.)
FADGI Application	Follow the guidance provided by the succession of EBU specifications as outlined on page 2. The ideal version indications would be as follows: • If neither a UMID nor loudness metadata is provided, then mark the file as version 0, using the value 0000h. • If a UMID is provided but not loudness metadata, then mark the file as version 1, using the value 0001h. • If both a UMID and loudness metadata are provided, then mark the file as version 2, using the value 0002h. Note that the EBU specification provides the flexibility to legally designate files as being in a higher version even if no values are provided for a UMID or loudness metadata. For example, a file may be marked as version 2 even if there is neither loudness information nor a UMID, or only one or the other. On the other hand, it is illegal to designate files as "lower version" when they do include an element value specified only for a higher version, e.g., it is wrong to mark a file as version 0 if it does have UMID or loudness information.
Mandatory/optional	Required by EBU specification
Values	0000h (version 0) 0001h (version 1) 0002h (version 2)
Example	0002h

UMID	
EBU Tech 3285	UMID 64 bytes containing a UMID (Unique Material Identifier) to the SPMTE 330M standard. If only a 32 byte basic UMID is used, the last 32 bytes should be set to zero. (The length of the UMID is given internally.) Note: The EBU intends to publish guidance on the use of the UMID in audio files. (Established in version 1.)

FADGI Application	Follow EBU specification. NOTE regarding the UMID identifier. The Working Group's understanding is that when file-reading software recognizes an instance of BWF version 1, it will look in the Version field to see if there is a UMID and report it. Readers that recognize an instance of version 0 will not look for a UMID. Meanwhile, no federal agencies are using UMID for audio files at this time and the Working Group has no particular recommendation regarding this identifier. NARA reports that their digitization systems produce version 1 files with the "slot" for the UMID left blank.
Char limit	64 bytes reserved for UMID
Mandatory/optional	Required by EBU specification

Reserved	
EBU Tech 3285	180 bytes reserved for extensions. If the Version field is set to 0001h, these 180 bytes must be set to a NULL (zero) value. (Version 0 established a 254-byte reservation; 64 bytes were removed in version 1 for the UMID; an additional 10 bytes were removed in version 2 for the five 2-byte loudness elements.)
FADGI Application	Follow EBU specification.
Char limit	190 bytes
Mandatory/optional	Required by EBU specification

I.B Recommended element for bext chunk

TimeReference	
EBU Tech 3285	This field contains the timecode of the sequence. It is a 64-bit value which contains the first sample count since midnight. The number of samples per second depends on the sample frequency which is defined in the field <nSamplesPerSec> from the <format chunk>. (Established in version 0.)
FADGI Application	The Working Group believes that TimeReference can be used in a manner that supports the management and synchronization of files that are parts of multitrack or multisegment works and plans to provide recommendations in the future.
Char limit	64-bit value
Mandatory/optional	Optional but recommended for complex recordings
Example	2374

I.C Optional elements for bext chunk

OriginationTime	
EBU Tech 3285	Eight ASCII characters containing the time of creation of the audio sequence. (Established in version 0.)
FADGI Application	This element contains the file creation time. This is understood to mean the local time in the timezone for the archival entity; the structure of the bext chunk does not permit ISO 8601 datetime indication, which unambiguously indicates date and time in terms of UTC (Coordinated Universal Time or Temps Universel Coordonné).
Char limit	8
Mandatory/optional	Optional
Values	ISO 8601 HH:MM:SS. Hour is defined from 00 to 23 (use leading zeroes if less than 10). Minute and second are defined from 00 to 59 (use leading zeroes if less than 10). The separator between the items is a colon [:] with no spaces. Note: the Working Group adheres to the ISO 8601 truncation convention: the hour should always be given but, if unknown, values for minutes and second are omitted. Thus the string length may be two, five, or eight. If the string entered is less than eight characters, end the data with a null character.
Example	01:45:25

CodingHistory	
EBU R98	Non-restricted ASCII characters, containing a collection of strings terminated by CR/LF. Each string contains a description of a coding process applied to the audio data. Each new coding application is required to add a new string with the appropriate information. This information must contain the type of sound (PCM or MPEG) with its specific parameters: PCM: mode (mono, stereo), size of the sample (8, 16 bits) and sample frequency: MPEG : sample frequency, bit-rate, layer (I or II) and the mode (mono, stereo, joint stereo or dual channel). It is recommended that the manufacturers of the coders provide an ASCII string for use in the coding history. (Established in version 0.)
FADGI Application	This element is designed to hold data on the digitizing process including signal chain specifics, sample rate and bit depth, and other elements. It is defined as a collection of strings, each presented on a separate line, containing a history of the coding processes applied to the file. The first line documents the analog source recording, the second line contains data on the capture process, the third line of data records information on the storage of the file. A new line is added when the coding history related to the file is changed. FADGI extends values beyond what is listed in EBU R98 to reflect current needs and tool capabilities: • sampling frequency: FADGI adds 96000, 176400, 192000, 384000, 768000 Hz • word length (also known as bit depth): FADGI adds 32 bits • mode: FADGI adds multitrack, multichannel, other **Notes on Mode** EBU R98 does not provide a definition for “mode” other than the defined values of “mono, stereo, dual-mono, joint-stereo”. Sound Directions (p. 41) states that

	mode corresponds to “sound field, such as mono, stereo, or dual-mono.” The FADGI glossary definition for sound field states it “Indicates the aural space arrangement of a mono, stereo, or multichannel sound recording, e.g., monaural, stereo, joint stereo, surround sound 5.1, etc. The term generally is used in a playback context, where a set of loudspeakers have been set up, e.g., for a surround sound presentation in a theater. The term may be compared to a visual field ("what you see before you"). Some multichannel sound formats include a channel map that defines the desired sound field” [such as left front loudspeaker, right rear loudspeaker].” This definition aligns with that for soundField in AES 57: AES standard for audio metadata - Audio object structures for preservation and restoration: “The <soundField> element shall be used to declare the overall soundstage to which the described audio object or region belongs. Valid values are mono, stereo, and surround.” To clarify the use of this field, **FADGI states that ‘mode’ is a synonym for ‘sound field’.** **Definitions for mode/sound field values:** **Mono: audio from a single recording source recorded onto a single channel or track** An item with true mono or monaural audio was originally recorded with a single recording source or pickup. This includes audio from a cylinder recording; a mono disc cutter head; full-track mono magnetic tape record head; and only one channel from a ½-Track or ¼-Track magnetic tape record head, for example. **Stereo: two-channel audio with different but related content in the left (01) and right (02) channels** Stereo describes the recording of a single instance or moment, where channels 1 and 2 present a “soundstage” or “soundfield” where the sound elements can be heard across a range of left and right. This includes audio that was recorded with a stereo disc cutter head with left and right channels or two-channel audio that was recorded with both channels of a 1/2-Track or 1/4-Track recording head. This can be, for example, an original recording using two microphones that captured a single performance OR a multitrack recording that was mixed down to stereo. **Dual-mono: any two-channel audio that is not stereo** Dual-mono is a catch-all mode for any two-channel audio that is not stereo (i.e. the audio doesn’t represent a coherent left/right “soundstage” or “soundfield”). For example, dual-mono can come from a ¼” open reel tape that has a mono program recorded on one side and another, different, mono program recorded on the other side. Dual-mono is also the mode applied to a mono (single) source that has been split, or duplicated, across a two-channel recording head. This results in two-channel audio with roughly the same content in both channels. **Joint-stereo:** Taking advantage of redundancies in audio coding resulting in the left and right stereo pair being coded in a single bitstream. joint-stereo refers to a compression technique where two stereo channels of audio frames are so perceptively similar that they would be encoded as a single channel or a single channel with some panning instruction rather than stereo. Methods for encoding include Intensity Stereo (combining the left and right audio channels at certain frequencies into mono) or Mid/Side Stereo (also known as matrix stereo or M/S stereo) where redundant content in left and right audio channels is combined into one Mid channel with non-redundant content stored in Side channel. See EBU TR49 Conversion Techniques For Multichannel Audio Formats (June

	2019) for more context on appropriate use. **Multitrack** A multitrack recording consists of multiple, separate streams of sound without defined locations in a sound field. During a typical music-recording session, individual instruments or voices will be recorded on separate tracks. When these multiple tracks are mixed at the end of the session, combinations of sounds are assigned places in a sound field for future playback. **Multichannel** A multichannel recording is a mixed and edited presentation in which multiple audio streams have been assigned locations in an expanded sound field ("left front loudspeaker, right rear loudspeaker") in order to provide “ambiance” or sound information that is happening somewhere other than from the left or right speaker. When implemented, use “multichannel” as the mode/sound field and use the “T=” to further describe the channel arrangement such as interleaved, surround, immersive, 5.1 (as described by EBU Recommendation ITU-R BS.775), Dolby Atmos and more. **Other** Used for instances where the configuration of the audio is not covered by another mode value option. When implemented, use “other” as the mode/sound field and use the “T=” for further description.
Char limit	Non-restricted
Mandatory/optional	Optional
Values	Each variable within a string is separated by a comma-space and each line should end with a carriage return and line feed. Summary of subelements: A=coding algorithm F=sampling frequency B=bit rate (only for MPEG) W=word length/bit depth M=mode/sound field T=free ASCII text string; contains no commas but semicolons may be used Detail on subelement syntax: A = Coding Algorithm <ANALOG, PCM, MPEG1L1, MPEG1L2, MPEG1L3, MPEG2L1, MPEG2L2, MPEG2L3> F=Sampling frequency <11000, 22050, 24000, 32000, 44100, 48000, 96000, 176400, 192000, 384000, 768000> Implied unit of measure [Hz] B (ONLY FOR MPEG ENCODING) = Bit-rate <any bit-rate allowed in MPEG 2 (ISO/IEC13818-3)>, Implied unit of measure [kbit/s per channel] W= Word Length/Bit Depth <8, 12, 14, 16, 18, 20, 22, 24, 32> Implied unit of measure [bits] M=Mode/Sound Field <mono, stereo, dual-mono, joint-stereo, multitrack, multichannel, streams >

	T=Text, free string <a free ASCII-text string for in house use. This string should contain no commas (ASCII 2Chex). Examples of the contents: ID-No; codec type; A/D type; track number for multitrack recordings, description of channel layout for multichannel audio, number and arrangement of streams>
Example Sound Directions	A=ANALOG,M=mono,T=Studer816; SN1007; 15 ips; open reel tape, A=PCM,F=96000,W=24,M=mono,T=Pyramix1; SN16986, A=PCM,F=96000,W=24,M=mono,T=Lynx; AES16; DIO, **(* see note below about EOL comma use)** Explanation: Line 1 reads: an analog, mono, open-reel tape played back on a Studer 816 tape machine with serial number 1007 at tape speed 15 ips. Line 2 reads: tape was digitized to PCM coding in mono mode at 96 kHz sampling frequency and 24 bits per sample on a Pyramix 1 DAW with serial number 16986. Line 3 reads: the audio was stored as a BWF file with PCM coding in mono mode at 96 kHz sampling frequency and 24 bits per sample using a Lynx AES16 digital input/output interface **NOTE:** These examples from the Sound Directions project include a comma (“,”) at the end of each line of text but the EOL comma is not included in EBU R98. FADGI is including the comma in this document to faithfully represent the Sound Directions example but FADGI does not require EOL commas.
Example	A=ANALOG,M=stereo,T=Otari MX5050; SN3690; 15 ips; open reel tape A=PCM,F=48000,W=18,M=stereo,T=NVision; NV1000; A/D
Example	A=Analogue, M=multitrack, T=8 track; 1” tape A=PCM, M=mono, T=track 1 of 8; captured by Merging Horus SN1334; Pyramix ver. 12.0.5 *** A=Analogue, M=multitrack, T=8 track; 1” tape A=PCM, M=mono, T=track 2 of 8; captured by Merging Horus SN1334; Pyramix

	ver. 12.0.5
Example	A=Analogue, M=multitrack, T=8 track; 1” tape A=PCM, M=multichannel, T=8 channel interleaved MBWF; captured by Merging Horus SN1334; Pyramix ver. 12.0.5
Example	A=Analogue, M=surround, T=12”; Quadraphonic; vinyl disc; 33.3 RPM A=PCM, M=multichannel, T=4 channel interleaved MBWF; captured by Merging Horus SN1334; Pyramix ver. 12.0.5

LoudnessValue	
EBU Tech 3285	A 16-bit signed integer, equal to *round* (100x the Inegrated Loudness Value of the file in LUFS). (Established in version 2.)
FADGI Application	None developed at this time.
Mandatory/optional	Optional
Values	*See table in the Appendix for treatment of Loudness parameters* For LoudnessValue, MaxTruePeakLevel, MaxMomentaryLoudness and MaxShortTermLoudness, the range of valid values is D8F1h to FFFFh (corresponding to the floating-point equivalent values of -99.99 to -0.01) and 0000h to 270Fh (0.00 to 99.99). The most significant bit of the 16-bit hexadecimal number is the sign bit; hence, values between 8000h and FFFFh represent negative numbers.

LoudnessRange	
EBU Tech 3285	A 16-bit signed integer, equal to *round* (100x the Integrated Loudness Range of the file in LU). (Established in version 2.)
FADGI Application	None developed at this time.
Mandatory/optional	Optional
Values	*See table in the Appendix for treatment of Loudness parameters* The range of valid values is 0000h to 270Fh (0.00 to 99.99). The most significant bit of the 16-bit hexadecimal number is the sign bit; hence, values between 8000h and FFFFh represent negative numbers. Therefore, if 7FFFh occurs, it is known that that particular parameter must be ignored. If any parameters are found to have values outside their valid ranges (not just 7FFFh) when reading the chunk then they shall be ignored too.

MaxTruePeakLevel	
EBU Tech 3285	A 16-bit signed integer, equal to round (100x the Maximum True Peak Value of the file in dBTP). (Established in version 2.)
FADGI Application	None developed at this time.
Mandatory/optional	Optional
Values	*See table in the Appendix for values* For LoudnessValue, MaxTruePeakLevel, MaxMomentaryLoudness and MaxShortTermLoudness, the range of valid values is D8F1h to FFFFh (corresponding to the floating-point equivalent values of -99.99 to -0.01) and 0000h to 270Fh (0.00 to 99.99). The most significant bit of the 16-bit hexadecimal number is the sign bit; hence, values between 8000h and FFFFh represent negative numbers.

MaxMomentaryLoudness	
EBU Tech 3285	A 16-bit signed integer, equal to *round* (100x the highest value of the Momentary Loudness Level of the file in LUFS). (Established in version 2.)
FADGI Application	None developed at this time.
Mandatory/optional	Optional
Values	*See table in the Appendix for treatment of Loudness parameters* For LoudnessValue, MaxTruePeakLevel, MaxMomentaryLoudness and MaxShortTermLoudness, the range of valid values is D8F1h to FFFFh (corresponding to the floating-point equivalent values of -99.99 to -0.01) and 0000h to 270Fh (0.00 to 99.99). The most significant bit of the 16-bit hexadecimal number is the sign bit; hence, values between 8000h and FFFFh represent negative numbers.

MaxShortTermLoudness	
EBU Tech 3285	A 16-bit signed integer, equal to *round* (100x the highest value of the Short-term Loudness Level of the file in LUFS). (Established in version 2.)
FADGI Application	None developed at this time.
Mandatory/optional	Optional

Values	*See table in the Appendix for treatment of Loudness parameters* For LoudnessValue, MaxTruePeakLevel, MaxMomentaryLoudness and MaxShortTermLoudness, the range of valid values is D8F1h to FFFFh (corresponding to the floating-point equivalent values of -99.99 to -0.01) and 0000h to 270Fh (0.00 to 99.99). The most significant bit of the 16-bit hexadecimal number is the sign bit; hence, values between 8000h and FFFFh represent negative numbers.

PART II. THE INFO CHUNK (“INFO list chunk”)

II.A Recommended element for info chunk

IARL	
RIFF Spec	Archival Location: Indicates where the subject of the file is archived.
FADGI Application	This element is an approximate equivalent to the Originator element in the bext chunk which records the entity responsible for the creation, maintenance, preservation of this digital item. If used, this element could repeat the data from the Originator element in the same structured format. In some applications, it is automatically carried over to the MP3 ID3 element "ArchivalLocation." Of all the info list chunk elements, this is the most important for Federal Agencies because it is vital to document the archival responsibility for a content item. Entity designations should be as specific as possible including a two- character county code to avoid the potential for conflict in the archiving entity’s name. The archival entity should be identified at the most specific level within the institution. Use a standard abbreviation of entity names such as those found in the Guide to Government Acronyms & Abbreviations. If an entity is not on this list, use a familiar abbreviation. Use the standard two-character ISO 3166 alpha 2 country code list.
Mandatory/optional	Strongly recommended (“required”) if using info list chunk
Values	[Country code]comma space[Entity name]
Example LC	US, LOC/RSS [RSS = Recorded Sound Section]
Example NARA	US, NARA

II.B Optional elements for info chunk

INAM	
RIFF Spec	Name: Stores the title of the subject of the file, such as, Seattle from Above.
FADGI Application	If used, this element should contain a *working title*. This phrase is intended to convey the Working Group’s understanding that metadata may be updated over time. Users who refer to elements like titles in an embedded block of metadata should understand that this may not represent the latest and best information, and they should be counseled to follow one of the identifiers in order to obtain later and/or better data.
Mandatory/optional	Optional
Values	Free text
Example LC	Symphony no. 3 in A minor, op. 56

Example LC	Interview with Bo Dollis at WWOZ, New Orleans, 1999-03-24 [title assigned by cataloger]
Example LC	Adventures in research. No. 587, Sawing off Manhattan Island [episode in a radio series]

ICMT	
RIFF Spec	Comment: Provides general comments about the file or the subject of the file. If the comment is several sentences long, end each sentence with a period. Do not include newline characters (such as CR/LF/EOL).
FADGI Application	If used, this element should repeat the principal and other identifiers from the OriginatorReference and/or Description element in the bext chunk. Each identifier will be labeled as to its origin or purpose using the “type” and “comment” qualifiers. Using the tag “local” in the "type" qualifier implies the entity identified in the bext Originator element. In some applications, this info list element may carry forward automatically to the MP3 ID3 tagged element "Comment." NOTE: The Working Group perceived value in the practices of labeling identifiers (see examples) but wished to leave this as an optional practice.
Mandatory/optional	Optional
Values	If labeled: Identifier [comma space] type [comma space] comment [semicolon-space if more than one identifier] If no labeling: Identifier
Example NARA	[Tagged identifier] 58979818, local, principal ID original filename
Example NARA	[Tagged identifier] 306-MUSA-9658B, local, RG-Series-Item Number
Example LC	[Two tagged identifiers, from LC, RSS] 1201566-2-1, local, system-generated number; RYI 6039, local, source location number
Example LC	[Tagged identifier] http://hdl.loc.gov/loc.mbrsmi/westhpp.2033, URL, principal ID Handle
Example California Digital Library	[Tagged identifier] ark.cdlib.org.org/ark:/13030/tf5p30086k, URL, ARK
Example LC	[Unlabeled identifier] http://lccn.loc.gov/mp76000002

ICRD	
RIFF Spec	Creation date: Specifies the date the subject of the file was created. List dates in year-month-day format, padding one-digit months and days with a zero on the left. For example, 1999-05-03 for May 3, 1999.
FADGI Application	This element is an equivalent to the OriginationDate in the bext chunk. If used, this element could repeat the information from the OriginationDate element in the bext chunk using the same structured format. In some applications, this info list element may carry forward automatically to MP3 ID3 tagged element "CreationDate."
Mandatory/optional	Optional
Values	ISO 8601 YYYY-MM-DD. Year is defined from 0000 to 9999; Month is defined from 01 to 12 (use leading zeroes if less than 10); Day is defined from 01 to 28, 29, 30 or 31 (use leading zeroes if less than 10). The separator between the items is a hyphen [-] with no spaces. **Note:** the Working Group adheres to the ISO 8601 truncation convention: the year should always be given but, if unknown, values for month and day are omitted. When the date is represented by a range, such as 2008--2009, double hyphens [--] may be used as the interval designator in place of a solidus (a.k.a. forward slash "/").
Example	2005-11-30

ICOP	
RIFF Spec	Copyright: Records the copyright information for the file. For example, Copyright Encyclopedia International 1991. If there are multiple copyrights, separate them by a semicolon followed by a space.
FADGI Application	Information about copyright and other restrictions (donor, privacy, etc.). Usage by federal agencies will often refer to the documentation of restrictions provided by other, non-embedded metadata. This element has no equivalent in the bext chunk. In some applications, this info list element may carry forward automatically to MP3 ID3 tagged element "Copyright." If used, this element may contain the information as known at the time of embedding. The Working Group understands that metadata may be updated over time. Users who refer to an embedded block of metadata should understand that this may not represent the latest and best information, and they should be counseled to follow one of the identifiers in order to obtain later and/or better data.
Mandatory/optional	Optional
Values	Free text. If there are multiple copyrights or other restriction statements, separate them by a semicolon followed by a space.
Example LC	Publication and other forms of distribution may be restricted. For details, contact the Recorded Sound Section of the Library of Congress. [Preferred wording from LC, RSS]
Example LC	(p) Rhino Records 2002.
Example LC	See Copyright Restriction Statement. [Used by LC, American Folklife Center]

Part III. ASSOCIATED DATA LIST CHUNK

The Associated Data List chunk provides labels, comments and other context for the Cue Chunk and individual cue points within the Cue Chunk.

The cue chunk (cue) is an optional, non-repeatable chunk in WAVE files that contains any number of Cue Points. A cue point is a specific point of special interest in the audio waveform data such as a change in speaker, start of a speech or vocal arrangement just to name a few examples. Cue points are sometimes referred to as flags or markers in digital audio applications. The cue chunk can be at any position in the file and should be ignored/not altered by programs which do not understand or recognize the chunk.

The cue chunk, when it occurs, contains one Chunk ID with the value of “cue ”, (*note the space after the word cue so that the chunk ID conforms to the 4 ASCII character rule*), a declaration of the chunk size (which is variable depending on the number of cue points listed within), a count of the number of cue points and finally, a list of the individual cue points in consecutive order.

Each cue point is further defined by a unique ID (which is used by other chunks in the file as a reference point – See adtl, plst and sample) as well as a variety of ways to determine the specific location of the cue point by using other data structures such as byte offset values to the data chunk and the block containing the sample.

Context for individual cue points is defined not in the cue chunk but in the Associated Data List Chunk (adtl) and its subchunks Label (labl), Note (note) and Labeled Text (ltxt). The key for associating each of these chunks to specific cue points is the unique ID for each cue point defined in the Cue Chunk.

If the cue Chunk is present, the adtl chunk is Strongly Recommended in order to provide content and context to the Cue Points. The labl chunk and note chunk describe one specific Cue Point whereas the Labeled Text (ltxt) chunk describes a range or section of audio data starting with a specific Cue Point and spanning across a declared sampling length. A Cue Point ID can be referenced by more than one adtl subchunk. In other words, a specific Cue Point can have a labl text label, a note text comment as well as a ltxt text description.

III.A Strongly Recommended element for Associated Data List chunk (adtl)

If the Working Group had the authority to do so, this would be "required" if implementing the Cue chunk.

List of Text Labels and Names	
RIFF Spec	The associated data list provides the ability to attach information like labels to sections of the waveform data stream.
FADGI Application	The adtl chunk defines which text labels and names which are associated with the cue points to provide each text label or name a position. The adtl chunk does not contain labels itself. Rather it declares where other data chunks contain this information. For WAVE files, the usual options are the Label Chunk (labl), Note Chunk (note) and Labeled Text Chunk (ltxt). Each adtl chunk has a chunk ID (“list”), a declaration of the chunk size (which is variable depending on the number and size of the subchunks), the chunk type (“adtl”) and a text list of the chunks that contain labels or other contextual information. (Label Chunk, Note Chunk and Labeled Text Chunk). This element is the list of adtl subchunks included in the file.
Char limit	Variable length
Mandatory/optional	Strongly Recommended if implementing Cue Chunk
Values	List as many as needed to indicate the chunk’s presence in the file: • labl = Label chunk • note = Note Chunk • ltxt = Labeled Text Chunk

III.B Recommended element for Label chunk (labl)

Text	
RIFF Spec	Specifies a NULL-terminated string containing a text label (for the 'labl' chunk) or comment text (for the 'note' chunk)
FADGI Application	The labl ‘text’ element is the **primary label** of the specific Cue Point and this information may be displayed next to markers, flags or cues in digital audio editors. The label can be of variable length but should be short and descriptive. More context about the label can be added in the ‘note’ chunk if needed. The labl text and note text elements are not linked directly to each other but rather both point back to the Cue Point ID. If there is a choice to use only one text element, use the labl text to provide a descriptive label and only use the note text element to provide more information or expand upon the labl text label. Each labl chunk comprises a chunk ID (“labl”), a declaration of the chunk size (which is variable depending on the size of the label text), the unique Cue Point ID from the Cue Chunk to which this label refers, and a text string of characters to define the label. This element is the text string of characters to define the label.
Char limit	Variable length
Mandatory/optional	Strongly Recommended if implementing Cue Chunk
Values	Free text string. ASCII.

III.C Optional element for Note chunk (note)

Text	
RIFF Spec	Specifies a NULL-terminated string containing a text label (for the 'labl' chunk) or comment text (for the 'note' chunk)
FADGI Application	The note ‘text’ element associates a **comment** to a specific Cue Point, either further explaining the labl text label or otherwise providing additional context. Like the labl text label, this information may be displayed next to markers, flags or cues in digital audio editors. The labl text and note text elements are not linked directly to each other but rather both point back to the Cue Point ID. If there is a choice to use only one text element, use the labl text to provide a descriptive label and only use the note text element to provide more information or expand upon the labl text label. Each note chunk comprises a chunk ID (“note”), a declaration of the chunk size (which is variable depending on the size of the comment text), the unique Cue Point ID from the Cue Chunk to which this comment refers, and a text string of characters to define the comment. This element is the text string of characters to define the comment.
Char limit	Variable length

Mandatory/optional	Recommended if implementing Cue Chunk in order to label the Cue Points.
Values	Free text.

III.D Recommended elements for Labeled Text chunk (ltxt)

Purpose	
RIFF Spec	Specifies the type or purpose of the ltxt text data. For example, Purpose can specify a FOURCC code like 'scrp' for script text or 'capt' for close-caption text.
FADGI Application	The purpose element works in concert with the ltxt text element with the purpose element defining the context of the information in the ltxt text element.
Char limit	4
Mandatory/optional	Recommended if implementing ltxt text element
Values	Choose one option from the list below. [1] • spea = speaker; to indicate a specific speaker name • envi= environment noises like mic feedback, laughter, paper rustling, echoes, etc • note = notes about the recording (ex: “possible cut in recording”) • tran = transcription • othr = other

Text	
RIFF Spec	The text is a null terminated string of characters. If the number of characters in the string is not even, padding must be appended to the string. The appended padding is not considered in the note chunk's chunk size field.
FADGI Application	This ltxt text element contains information for a range of cue points. The ltxt purpose element defines the context for the ltxt text data
Char limit	Variable
Mandatory/optional	Recommended if implementing ltxt purpose element
Values	Free text string. ASCII

[1] This controlled vocabulary was inspired by the “SENT” structure (speaker, environment, note, transcription) for IIIF Annotation Layers developed by Kylie Warkentin at the University of Texas at Austin in the AudiAnnotate project.

III.E Optional elements for Labeled Text chunk (ltxt)

Country	
RIFF Spec	Specifies the country code used for file elements. Defined by values in the CSET chunk.
FADGI Application	Country location information used in the text strings in the internal file metadata. This element does not document the country of the entity responsible for the creation, maintenance, preservation of this digital item. This information is recorded in the Originator (bext) and IARL (list) fields. FADGI recognizes that the values for country codes, languages and dialects defined in the CSET chunk are outdated because the RIFF specification in which they are defined was last updated in 1991. To ensure maximum compatibility, use the CSET codes if the appropriate values are defined in the CSET chunk, otherwise use ISO standard code lists: ISO 3166-1 code for Country or ISO 639-1 for Language
Mandatory/optional	Optional
Values	For maximum compatibility, use country codes as defined in the RIFF CSET chunk. If the CSET chunk is not present, or if this field has value zero, assume USA (country code 001). If the country is not defined in the CSET chunk, use the ISO 3166-1 code for Country.

Language	
RIFF Spec	Specify the language used for file elements. Defined by values in the CSET chunk.
FADGI Application	Specifies the language used in the text strings in the internal file metadata. This element does not define the language in the audio stream content. FADGI recognizes that the values for country codes, languages and dialects defined in the CSET chunk are outdated because the RIFF specification in which they are defined was last updated in 1991. To ensure maximum compatibility, use the CSET codes if the appropriate values are defined in the CSET chunk, otherwise use ISO standard code lists: ISO 3166-1 code for Country or ISO 639-1 for Language
Mandatory/optional	Optional
Values	For maximum compatibility, use the Language Codes in RIFF CSET chunk. If the CSET chunk is not present, or if these fields have value zero, assume US English (language code 9, dialect code 1). If the language is not defined in the CSET chunk, use the ISO 639-1 values for language. Note: ISO 639-1 does not provide an option for dialects to indicate regionality. If the ISO 639-1 values are used for Language, the Dialect element value defaults to 0 (zero). However the combined values of Country and Language can provide more specific location information for the language.

Dialect	
RIFF Spec	Specify the language dialect used for file elements. Defined by values in the CSET chunk.
FADGI Application	Specifies the dialect of the language used in the text strings in the internal file metadata. This element does not define the dialect of the language in the audio stream content. If the Language element uses CSET values, then the Dialect element may also be present and, if present, must use CSET values as well. The ISO 639-1 code for Language do not include an option for Dialect. If the Language element uses ISO 639-1 values, then the Dialect element will default to 0 (zero). The combined values of the ISO 639-1 Language code and the Country code (either CSET or ISO 3166-1 values) can provide additional information about the language regionality.
Mandatory/optional	Optional
Values	Language and Dialect Codes in RIFF CSET chunk. If the CSET chunk is not present, or if these fields have value zero, assume US English (language code 9, dialect code 1). Note: ISO 639-1 does not provide an option for dialects. If the ISO 639-1 values are used for Language, the Dialect element value defaults to 0 (zero).

Code Page	
RIFF Spec	Specifies the code page for the text.
FADGI Application	Same. The code page defines the character encoding for the text strings in the internal file metadata such as INFO list and bext chunks. The values for the code page are defined in the CSET chunk.
Mandatory/optional	Optional
Values	If the CSET chunk is not present, or if this field has value zero, assume standard ISO 8859/1 code page (identical to code page 1004 without code points defined in hex columns 0, 1, 8, and 9).

Part IV. APPENDIX: TREATMENT OF LOUDNESS PARAMETERS

The following text is from section 2.4 in the version 2 of the Broadcast WAVE specification (EBU Tech 3285 v2, May 2011; http://tech.ebu.ch/docs/tech/tech3285.pdf) . Additional general information is offered in the related EBU document Loudness normalisation and permitted maximum level of audio signals (EBU R 128, http://tech.ebu.ch/docs/r/r128.pdf, August 2011; http://tech.ebu.ch/docs/r/r128.pdf; both URLs accessed April 17, 2012). A key work on loudness metrics has been published by the International Telecommunications Union- Radiocomunication Sector (ITU-R): Algorithms to measure audio programme loudness and true-peak audio level (ITU-R BS.1770-2, March 2011; http://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1770-2-201103-I!!PDF-E.pdf).

Treatment of Loudness Parameters

The loudness parameters are represented by integers, but they preserve a precision of two decimal places by being multiplied by 100 before being rounded. The rounding function which shall be used is defined as follows:

integer representation = integer part of (x + sgn(x) · 0.5) where x is the value to be represented, multiplied by 100

and where sgn() is the signum operator. sgn(x) = -1 if x < 0, 0 if x = 0, 1 if x > 0.

This rounding method is commonly referred to as “round to nearest, ties away from zero” because where the fractional part of the number is 5 (midway between integers), the rounding is up for positive numbers and down for negative numbers.

Examples

Negative numbers: Float value	**Calculation**	**Value carried in BWF (decimal/ hexadecimal)**
-22.644	integer[(-22.644 x 100) + sgn(-22.644 x 100) · 0.5]	-2264/ F728h
-22.645	integer[(-22.645 x 100) + sgn(-22.645 x 100) · 0.5]	-2265/ F727h
-22.646	integer[(-22.646 x 100) + sgn(-22.646 x 100) · 0.5]	-2265/ F727h
Positive numbers: Float value	**Calculation**	**Value carried in BWF (decimal/ hexadecimal)**
12.764	integer[(12.764 x 100) + sgn(12.764 x 100) · 0.5]	1276/ 04FCh
12.765	integer[(12.765 x 100) + sgn(12.765 x 100) · 0.5]	1277/ 04FDh
12.766	integer[(12.766 x 100) + sgn(12.766 x 100) · 0.5]	1277/ 04FDh

Part V. SPECIFICATIONS AND REFERENCES

V.A Specifications relevant to the preceding recommendations.

European Broadcast Union. *Loudness normalization and permitted maximum level of audio signals.* August 2020. Online: http://tech.ebu.ch/docs/r/r128.pdf, accessed: March 2, 2021.

European Broadcast Union. *Technical Recommendation R98-1999. Format for the <CodingHistory> field in Broadcast WAVE Format files, BWF.* 1999. Online: http://tech.ebu.ch/docs/r/r098.pdf, accessed: March 2, 2021.

European Broadcast Union. *Specification for the Broadcast WAVE Format (BWF): A format for audio data files in broadcasting. Version 2. May 2011.* Online: http://tech.ebu.ch/docs/tech/tech3285.pdf, accessed: March 2, 2021

European Broadcast Union. *BWF: A format for audio data files in broadcasting – Supplement 2 – Capturing Report. July 2001. Tech3285-2.* Online: https://tech.ebu.ch/docs/tech/tech3285s2.pdf, accessed: March 2, 2021

IBM Corporation and Microsoft Corporation. *Multimedia Programming Interface and Data Specifications 1.0.* Online: https://www.aelius.com/njh/wavemetatools/doc/riffmci.pdf, accessed: March 2, 2021.

International Organization for Standardization. *ISO 3166-1 Country Codes*. Full codes available online through Wikipedia: https://en.wikipedia.org/wiki/ISO_3166-1_alpha-2, accessed: March 2, 2021.

International Organization for Standardization. *ISO 639-1 Language Codes*. Full codes available online through Wikipedia: https://en.wikipedia.org/wiki/List_of_ISO_639-1_codes, accessed: March 2, 2021.

Warkentin, Kylie. *Improving Gentle Transcription Using SENT Metadata Structure*. Part of University of Texas at Austin in the AudiAnnotate project. Online: https://kywark.github.io/gentle-improvement/, accessed: March 2, 2021

V.B Other References

Advanced Television Systems Committee. *ATSC Recommended Practice: Techniques for Establishing and Maintaining Audio Loudness for Digital Television*, Document A/85:2013. Online: https://www.atsc.org/atsc-documents/a85-techniques-for-establishing-and-maintaining-audio-loudness-for-digital-television/, accessed March 2, 2021.

Audio Engineering Society. *AES57-2011 (r2017): AES standard for audio metadata - Audio object structures for preservation and restoration*. Online: https://www.aes.org/publications/standards/search.cfm?docID=84. accessed: March 2, 2021.

Association of Recorded Sound Collections. *ARSC Guide to Audio Preservation*. 2015. Online: https://www.clir.org/wp-content/uploads/sites/6/pub164.pdf, accessed: March

2, 2021

Casey, M. and B. Gordon. 2007. *Sound Directions: Best Practices For Audio Preservation.* Trustees of Indiana University and President and Fellows of Harvard University. Online: http://www.dlib.indiana.edu/projects/sounddirections/papersPresent/sd_bp_07.pdf, accessed: March 2, 2021.

Camerer, Florian. "Overview of the EBU Loudness Recommendation R 128." *SMPTE Motion Imaging Journal.* July/August 2011.

European Broadcast Union. *MBWF / RF64: An extended File Format for Audio. July 2009. Tech3306.* Online: https://tech.ebu.ch/docs/tech/tech3306v1_1.pdf, accessed: March 2, 2021

European Broadcast Union. *Conversion techniques or multichannel audio formats. June 2019. TR049.* Online: https://tech.ebu.ch/docs/techreports/tr049.pdf, accessed: March 2, 2021

European Broadcast Union. *Multichannel use of the BWF audio file format (MBWF). R1111.* Online: https://tech.ebu.ch/docs/r/r111.pdf, accessed: March 5, 2021

European Broadcast Union Project Group P/MCA (Multichannel Audio). *EBU's multichannel audio activities..* Online: https://tech.ebu.ch/docs/techreview/trev_292-multichannel_audio.pdf, accessed: April 26, 2021

Harvard University Library. *Administrative Metadata for Digital Audio Files.* Online through Internet Archive: https://web.archive.org/web/20160123041743/http://hul.harvard.edu/ois/systems/drs/audiometadata.pdf, accessed: March 2, 2021.

IASA Technical Committee. *Guidelines on the Production and Preservation of Digital Audio Objects*, ed. by Kevin Bradley. Second edition 2009. Standards, Recommended Practices and Strategies, IASA-TC 04. Online: www.iasa-web.org/tc04/audio-preservation, accessed March 2, 2021

Indiana University Archives of Traditional Music. *An Introduction to Technical and Structural Metadata Elements in ATMC.* Online: http://www.dlib.indiana.edu/projects/sounddirections/papersPresent/sd_app1_v1.pdf, accessed: March 2, 2021.

International Telecommunications Union. *Multichannel stereophonic sound system with and without accompanying picture. Recommendation ITU-R BS.775-3 (08/2012)*. Online: https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.775-3-201208-I!!PDF-E.pdf, accessed March 2, 2021.

Smith, Kelly L. *GovSpeak: A Guide to Government Acronyms & Abbreviations.* Online: http://libguides.ucsd.edu/govspeak, accessed: March 2, 2021.