SWGDE Best Practices for Digital Audio Authentication

Intellectual Property:
Unauthorized use of the SWGDE logo or documents without written permission from SWGDE is a violation of our intellectual property rights.

Individuals may not misstate and/or over represent duties and responsibilities of SWGDE work. This includes claiming oneself as a contributing member without actively participating in SWGDE meetings; claiming oneself as an officer of SWGDE without serving as such; claiming sole authorship of a document; use the SWGDE logo on any material and/or curriculum vitae.

Any mention of specific products within SWGDE documents is for informational purposes only; it does not imply a recommendation or endorsement by SWGDE.

SWGDE Best Practices for Digital Audio Authentication

Table of Contents

1. Scope & Purpose

1.1 Introduction

As defined in *SWGDE/SWGIT Digital & Multimedia Evidence Glossary* [1], authentication is the process of substantiating that the asserted provenance of data is true. As defined in *SWGDE Best Practices for Forensic Audio* [2], an audio authentication examination seeks to determine if a recording is consistent with the manner in which it is alleged to have been produced.

The term "authentication" is often used in a legal context to describe the establishment of a proper legal foundation for the admission of a recording as evidence into a judicial proceeding. This is generally accomplished by a party involved in the events recorded or involved in the recording process affirming that the events heard during playback of the recording are consistent with that party's recollection of the events as they transpired. When this is contested or cannot be accomplished, a scientific analysis may be conducted to test disputed claims.

1.2 Scope

The purpose of this document is to provide the background, technical considerations, and potential criteria upon which to conduct forensic authentication examinations of digital audio when its provenance and/or integrity is in question.

1.3 Limitations

This document does not address the authentication of analog audio recordings/media nor digital tape, or non-file based media.

This document is not intended to be a training manual or a specific operating procedure. This document is not all inclusive and does not contain information relative to specific commercial products. If dealing with technology outside your area of expertise, consult with an appropriate specialist. For recommendations on forensic audio training, refer to *SWGDE Core Competencies for Forensic Audio* [3] and *SWGDE/SWGIT Guidelines and Recommendations for Training in Digital and Multimedia Evidence* [4]. For recommendations on developing standard operating procedures refer to *SWGDE/SWGIT Recommended Guidelines for Developing Standard Operating Procedures* [5].

2. Terminology

- **byte offset**, *n*—the location or index of a data object (such as an alphanumeric character string) in a file or volume with respect to the first byte of data at location 0.
- **container**, *n*—the supporting data and metadata contained in a file which is accessed by the host system to enable playback.
- **content**, *n*—the binary data making up the audio portion of the file and the acoustic events it represents.
- **encode,** *v*—the process of quantizing analog data into a finite bit-stream, often involves the application of an algorithm designed to reduce the amount of data for storage or transmission.

- **first removable**, *adj*—either an original or duplicate recording that is transferrable, can be hashed, is write-protectable, or the physical media itself.
- **original recording,** *n*—the first manifestation of sound in a recoverable stored format.
  - Discussion—Not every original recording is recoverable on its storage device. The first removable instance of the original data should be protected from modification (write-blocking, over-record protection, etc.) and hashed if possible. Data integrity preservation is key to this definition. See *first removable*.
- **provenance**, *n*—the true history of a physical item and the recordings stored or encoded on or within it.
- **transcode,** *v*—to convert a multimedia file from one encoding scheme to another.

3. Pre-Examination

3.1 Clarifying the Request

Determine what specific questions the submitter is trying to answer and help clarify the request. Care should be taken to avoid the effects of potentially biasing information shared by the submitter. Factors to inquire about may include:

- Is the submitted exhibit the purported original recording? If not, request to obtain it. Please refer to Section 3 of [2] and the accompanying Appendix I Decision Tree for best practices related to evidence retrieval and retrieval methods.
  - There are some circumstances where the original, as defined by the first manifestation of recorded sound, is not obtainable or no longer exists. In these circumstances, a verified duplicate of the original media may be used for examination.
  - Transcoded or re-encoded versions of an original recording are not original recordings. A decoded PCM waveform copy of an encoded original may still contain useful information relevant to authenticity analysis, but file format analysis regarding the original format becomes impossible.
- What is the chain of custody of the purported original recording?
- What is the date and time the recording was purported to have been made?
- What is the purported original recording system, device, and storage media?
  - Request makes, models, and serial numbers, including microphones used. Other technical details, such as source of power, may be useful.
- Is this equipment used to make the recording accessible for testing?
- What was the physical location in which the recording was purportedly made and the description of its acoustic environment?
- Which parties were allegedly present when the recording was purportedly made?
- What are the specific questions or issues raised related to the evidence recording's authenticity?

- What file and audio encoding formats are supported by the purported original recording system?
  - What audio encoding formats can the device encode?
  - What file formats can the device produce/export?
  - What removable physical media can be used in the device?
  - Can the device transmit or receive electronically?
  - Determine if there are embedded security features?
- Any other relevant questions that derive from the discussion with the submitter.

Determine if more items are required to perform a complete exam, such as recording equipment and exemplar data or recordings. Access to the claimed recording location may also be necessary.

3.2 Assess the Request

Based on answers to the questions above, determine if an "authentication" exam is suitable to answer the submitter's questions. If not, work with submitter to modify the request as needed.

3.3 Define the Test Plan

Based on reported assertions related to the evidence, it may be useful to test those assertions as hypotheses during the examination.

If it is necessary to test an evidence device (e.g., the digital audio recorder purported to have recorded the evidence that does not have removable media), obtain permission from the necessary party(s) to do so and warn that this testing risks changing the state of that device.

Once the test plan has been established, the scope of the exam, as understood by the examiner, should be communicated and confirmed with the submitter.

4. Examination

4.1 Summary

The digital audio authenticity examination is a process which leads to a conclusion based on the interpretation of testing results, and shall be comprised of a clearly defined set of analyses. As with any scientific examination, the process shall be systematic, objective, and repeatable. Tests results should be reproducible. The process should be designed in a way that pays close attention to cognitive bias, its sources and influence, and in a way that mitigates its effects.

Use only scientifically valid peer-reviewed methods.

Analyses can be categorized as global and local. Global analyses are conducted upon the file as a whole and produce results relevant to a recording's authenticity without regard to specific areas or portions of a recording. Local analyses are conducted upon specific areas or portions of a recording and produce results relevant to a recording's authenticity. These analysis methods must be peer-reviewed. The tools which implement these methods along with the examiner's use and interpretation must be scientifically valid.

Local and global analyses can be classified as *observed* or *measurement based*. Format and structure, including counting (e.g. zero amplitude samples at the beginning/end of a file and quantization levels) are categorized as observational analysis. Measurement uncertainty or measurement error applies to measurement based analyses only.

Not all tests will be applicable in each case, nor shall any single analysis be relied upon solely. All applicable tests available to the examiner shall be conducted. By conducting multiple tests, the results from each can be cross-verified, thereby increasing confidence in the conclusion. [6]

During the examination, the test plan, procedures, data, and results shall be documented as well as the decision criteria used to reach the conclusion. Refer to Section 7.1 of [2] for more information.

See *Appendix A* for a sample case involving a commercially available digital audio recorder.

4.2 Assessment

Assess and document technical characteristics of the recording, which may include the following:

- container format, encoding scheme;
- channel configuration;
- bit rate, sample rate;
- duration;
- encoded date and time.

4.3 Critical Listening

Critical aural review of the submitted recording, without applying any processing, can yield information regarding areas to direct subsequent analyses. In addition to a preliminary overview, attention must be paid to voices, acoustic events, background noise changes, uncharacteristic noises that may indicate equipment malfunction and possible record and edit events, and any other areas of specific interest. To focus attention on each of these factors individually will require repeated playback [7]. Repeated aural review of the unprocessed and processed versions of the submitted recording may be necessary throughout the examination.

4.4 Global Analyses

The following is a non-exhaustive list of global analyses that can be applied.

4.4.1 File Format Analysis

This is a type of observational analysis in which an examiner uses applications capable of displaying multimedia file metadata, document metadata fields and their values.

Metadata could include: recorder make/model, serial number, settings applied during the recording, date and time of recording, recording length, and user data such as investigator or case number.

The choice of metadata application should be based on testing with known data prior to use in cases. In most situations, it is wise to use multiple metadata applications, or to confirm findings from one application with another because some applications extract more metadata results than others.

4.4.1.1 Tests that are applicable to proprietary recording systems, devices, and files

- If a proprietary recording system has built-in security or validation features, document the state of those features. Obtain documentation from the manufacturer of the embedded security features and include in case notes.
- If a recording system uses a file naming scheme, document the filename and its relationship to the naming scheme.
- If the evidence is in a proprietary file format, research the proprietary system to determine:
  - Are specific or proprietary applications or codecs required on the lab system to decode or play back the audio signal?
  - What is the impact of transcoding the file into a common format? Will metadata be preserved or changed?
  - How is the format structured?
  - Are there non-perceptual data streams or administrative data present (e.g., timecode, data block numbering) within the audio file or in separate files produced by the proprietary system as part of the recording process?
  - Are there unrelated files present?
  - Are other file remnants present?
  - Can recordings and administrative files deleted from the device be recovered?
  - Will proprietary software alter the data files in any way?
  - How does the proprietary system interface with the laboratory system and tools?

Other analyses may require that proprietary files be exported to a common format or captured during live playback.

4.4.2 File Structure Analysis

The file structure, including header metadata, recorded content, metadata multiplexed with the recorded content, and possible footer data of the submitted recording are observational analyses and should be compared with exemplar recordings made from submitted recorders or other test recordings. If the purported original recording device is not available or is otherwise unable to be utilized, use the same make/model of device. Document these structures to include both hex and American Standard Code for Information Interchange (ASCII) fields and their byte offsets.

File formats from proprietary recording systems and files encoded with proprietary codecs, such as the Digital Speech Standard (DSS), may be encountered. Analysis of these formats requires special consideration and may reveal:

- structure of recorded content in pre-defined data blocks;

- embedded checksum system within the recording structure;
- playback only available on manufacturer's proprietary software;
- a separate recording log file accompanying multimedia files;
- inclusion of proprietary playback software upon export.

4.4.2.1 Open source formats

Open source formats, such as Windows Media Audio (WMA), Moving Picture Experts Group layer-3 Audio (MP3), Windows Wave (WAV), Advanced Audio Coding (AAC), etc., adhere to commonly accepted specifications, which are compatible with most playback and editing software. The possibility exists that a recording may be altered using an external editor and placed back onto a device prior to submission and extraction. Audio editing software applications may modify a file's metadata fields, chunks, etc. and change their order, alter their information, add new ones, and/or delete them entirely. Therefore, when encountering these formats, it should be mandatory to supply the purportedly original file and perform digital data analysis on the metadata, and compare it to exemplars [8], [9], [10], [11], [12], [13].

Metadata may include recording make/model, recorder settings, recording start and stop dates/times, recording length, codec. Audio editing software applications often modify a file's metadata fields, chunks, etc. and change their order, alter their information, add new ones, and/or delete them entirely.

4.4.3 Quantization Level (QL) / Bit Depth Analysis

This method allows analysis of a signal's native bit depth to reveal consistencies with technical characteristics of the purported recorder, possible transcoding or signal processing, the bit depth trajectory of a signal, and to detect traces of local interpolation. [14] Quantization level analyses can be observational or measurement based. Measurement uncertainty in QL analysis is a function of bit depth and quantization error.

4.4.4 DC Offset

Direct current (DC) offset (the average direct current) occurs when one or more components in a recording system (such as microphones, microphone preamplifiers, interconnections, or recording circuitry) add DC voltages to a recorded audio signal resulting in a waveform that is not centered on the midline or x-axis [15].

Because DC offset is a result of a recorder's recording chain, measurements of a recording's DC signal may be used to locate indications of editing and, in certain circumstances, DC analysis can be used to test a claimed recorder given the same recording chain and environment. In this pursuit, it is necessary to conduct tests using the evidence recorder with the same settings, environment, and same/similar transmission channel characteristics (microphone). More precise results will be achieved with longer recordings (both evidence and exemplar) yielding more data [16].

The DC offset can be measured globally to provide an average value of an entire recording or in time window segments (locally) providing a vector of DC values for the recording. The global average DC offset alone is not typically enough for a conclusive examination.

Caution should be used when interpreting DC measurements made from data in which the signal is saturated (clipped) or numerically zero, as these conditions will affect the measured result but may be due to other factors such as post-processing. Measurement uncertainty in DC offset analysis is a function of the windowing length and quantization error as well.

Also, exemplar recordings should be made with all appropriate recorder settings.

4.4.4.1 Special considerations

- The recorded DC offset may change with recording configuration: internal/external mic/line input, recorder settings.
- The DC offset may also vary because of other factors, including a change in environment or location.

4.4.4.2 During a DC analysis test, the examiner seeks to answer various questions, including the following:

- Is the DC offset consistent throughout the questioned recording?
- Is the DC offset consistent between a known and the questioned recording?
- Is the DC offset consistent between channels?
- Is the DC offset difference between channels consistent between a known and the questioned recordings?

4.4.5 Long Term Spectral Analyses

Spectral analyses are measurement based. Their measurement uncertainty depends on sampling frequency, bit depth, windowing function, and FFT order.

4.4.5.1 Long Term Average Spectrum

The long term average spectrum (LTAS) of a signal is its plot of power in decibels (dB) as a function of frequency averaged from predefined fast Fourier transform (FFT) time windows. Inherently it is a function of the digital recorder's sample rate, but it may also be modified by the encoding algorithm. Therefore, characteristics of the LTAS can be observed to verify recorder settings and provenance.

4.4.5.2 Sorted Spectrum and Differentiated Sorted Spectrum

Long-term average sorted spectrum (LTASS) is LTAS sorted on descending order. Sorting the first derivative of the LTAS can detect traces of recompression in a first generation constant bit rate (CBR) compressed signal. [17]

4.4.6 Compression Level Analysis (CLA)

Compression level analysis is measurement based. The measurement uncertainty depends on sampling frequency, bit depth, windowing function, and FFT order. Because digitally sampled acoustic signals are essentially random, original uncompressed audio files have lower correlation between neighboring samples in the time domain. In the frequency domain, the lossy compressed algorithms have lower correlation. Conversely, some compressed audio files exhibit sample-to-sample correlation in the form of periodicity as a result of perceptual encoding. In understanding and applying these principles, one can analyze the compression level of an audio file to determine its provenance and, specifically, whether it is consistent with an original audio file or one that has been recompressed through a transcoding or re-encoding process. [18] [19]

4.4.7 Modified Discrete Cosine Transform (MDCT) Analysis

Lossy compression history assessment through MDCT coefficients is a measurement based analysis showing whether the evidence signal is consistent with an exemplary recording, or indicating traces of previous lossy recompression and edits. [20] [21] The MDCT measurement uncertainty depends on the lossy compression algorithm and bit rate.

4.4.8 Electric Network Frequency (ENF) Analysis

The ENF of a power grid can be induced in some recordings leaving a trace signal. The analysis and application of ENF can be useful in some circumstances. [22] [23] ENF is observational analysis when used to determine the type and number of ENF components. It is measurement based when used to extract and compare the evidence against a known database. Measurement uncertainty in ENF analysis depends on signal-to-noise ratio, distortion (e.g. clipping, EMF interference), sampling frequency, bit depth, and ENF extraction method.

ENF measurements can also be useful for local analyses to detect discontinuities.

4.5 Local Analyses

The following is non-exhaustive list of local analyses that can be applied.

4.5.1 Waveform Analysis

The waveform of an audio signal displays the relationship between time and amplitude of the acoustic information recorded, therefore allowing the determination of relative temporal and amplitude characteristics of a digital file ranging from a single sample point to an entire recording. Phenomena heard during critical listening can be visually observed as a waveform, such as dropouts, clipping, or other amplitude related events.

4.5.2 Signal Power

The power of an audio signal can be computed on short time frames, displayed as the trajectory of the power over the entire signal, and used to reveal possible missing information of the acoustic signal due to mechanical failure, compression artifacts, or insertion of silence. Signal power is observational and measurement based. Measurement uncertainty depends on distortion (e.g. clipping) and windowing length.

4.5.3 Spectrographic Analysis

Spectrograms display audio content in a frequency vs. time representation and may be used to gain an overall impression of the recording and to view specific local events. Examples of observable events include characteristics of digital aliasing, sample band-pass filtering, questioned signals, background sounds, convolution and transmission characteristics, and power line frequency components. Spectrographic analyses are observable and measurement based. Their measurement uncertainty depends on sampling frequency, bit depth, windowing function and length, and FFT order. [6] [17]

Spectrographic analysis can be used to plot an entire recording or locally on a targeted area.

4.5.4 Butt-Splice Detection

Butt-splice detection functions, such as application of the first derivative on a signal's amplitude, show the rate of change from one sample to the next, allowing for the detection of some crude edits. [24] Butt-splice detection is both observable and measurement based. Its measurement uncertainty depends on sampling frequency, bit depth, and windowing function and length.

4.6 Device Verification

Some analyses can be very useful in corroborating evidence with the purported recorder or attributing it to an unknown recorder. The following analyses can be useful in this way: Structure, Format, Long Term Average Spectrum, Sorted Spectrum, Differentiated Sorted Spectrum, DC Offset, and ENF.

4.7 Generation of sample recordings and exemplars

Either from the obtained evidence recorder or from test recorders of the same manufacturer and model, test recordings can be made and compared against the evidence recording. In this way, the evidence can be assessed as to its consistency with a known, authentic file.

Test recordings can also be obtained from a contributor that has prepared them or are available from the recorder's memory.

If making test recordings on an evidence device, this may change the state of the device and will change the state of non-removable media. If necessary and possible, produce a verified bit-stream copy of the non-removable memory prior to the testing.

5. Reporting Conclusions

5.1 Results Interpretation

Cross-verify the results of the analyses conducted. Scientifically interpret the results and provide technical explanation(s) in an impartial manner.

5.2 Conclusions

Authenticity conclusions should factually state the results of the scientific examinations, and should be thoroughly supported by the analyses conducted. The analysis conclusions should not be over or understated. Findings should be reported with regard to the analyses conducted.

It is possible for the results of an examination to be:

- Consistent with an original
- Inconclusive
- Inconsistent with an original

Additionally, discontinuities and alterations within the questioned recording should be included in the report with an explanation of the cause. Examples would include recording pauses, voice activation discontinuities, etc.

The findings can be communicated to the contributor and if deemed necessary, a written report prepared.

5.2.1 No scientific inquiry, including those in forensics, produce a result of absolute certainty. Therefore, conclusions in digital audio examinations related to an audio recording's authenticity should not be stated in terms of absolutes. Language implying 100% certainty should be avoided unless speaking about known alterations or deletions [17].

5.3 Report Components

Details typically included in a report regarding the authenticity of a recording will vary based on the analyses conducted. *ASTM Practice E620 for Reporting Opinions of Scientific or Technical Experts* [25] should be followed.

The following items are considered necessary:

- Description of the evidence and its chain of custody.
- Technical details of the submitted evidence and analyses conducted.
- Data and reasoning relied upon to come to conclusion(s).

Other relevant details may include assertions related to the evidence's authenticity and examination findings related to those assertions.

6. References

[1] Scientific Working Group on Digital Evidence, "SWGDE Digital & Multimedia Evidence Glossary". [Online]. https://www.swgde.org/documents

[2] Scientific Working Group on Digital Evidence, "SWGDE Best Practices for Forensic Audio". [Online]. https://www.swgde.org/documents

[3] Scientific Working Group on Digital Evidence, "SWGDE Core Competencies for Forensic Audio". [Online]. https://www.swgde.org/documents

[4] Scientific Working Group on Digital Evidence and Scientific Working Group on Imaging Technology, "SWGDE/SWGIT Guidelines & Recommendations for Training in Digital & Multimedia Evidence". [Online]. https://www.swgde.org/documents

[5] Scientific Working Group on Digital Evidence, "SWGDE Model Standard Operation Procedures for Computer Forensics". [Online]. https://www.swgde.org/documents

[6] B. Koenig and D. Lacey, "Forensic Authentication of Digital Audio Recordings," *Journal of the Audio Engineering Society*, vol. 57, no. 9, pp. 662-695, 2009.

[7] D. L. Rappaport, "Establishing a Standard for Digital Audio Authenticity: A Critical Analysis of Tools, Methodologies, and Challenges," Denver, 2012.

[8] B. Koenig and D. Lacey, "Forensic Authenticity Analyses of the Metadata in Re-Encoded WAV Files," in *AES 54th International Conference*, London, England, June 2014.

[9] B. Koenig, D. Lacey, and C. Reimond, "Selected Characteristics of MP3 Files Re-Encoded With Audio Editing Software," *Journal of Forensic Identification*, vol. 64, no. 3, pp. 304-321, May/June 2014.

[10] B. Koenig and D. Lacey, "Forensic Authenticity Analyses of the Header Data in Re-Encoded WMA Files From Small Olympus Audio Recorders," *Journal of the Audio Engineering Society*, vol. 60, no. 4, pp. 255-265, April 2012.

[11] C. Grigoras and J. M. Smith, "Large Scale Test of Digital Audio File Structrure and Format for Forensic Analysis," in *AES International Conference on Audio Forensics*, Arlington, VA, 2017.

[12] J. M. Smith, D. Lacey, B. Koenig, and C. Grigoras, "Triage Approach for the Forensic Analysis of Apple iOS Audio Files Recorded Using the "Voice Memos" App," in *AES International Conference on Audio Forensics*, Arlington, VA, 2017.

[13] B. Koenig and D. Lacey, "Forensic Authenticity Analyses of the Metadata in Re-Encoded iPhone M4A Files," in *AES International Conference on Audio Forensics*, Arlington, VA, 2017.

[14] C. Grigoras and J. Smith, "Quantization Level Analysis for Forensic Media Authentication," in *54th International Conference of the Audio Engineering Society*, London, 2014.

[15] B. Koenig, D. Lacey, C. Grigoras, S. Price, and J. Smith, "Evaluation of the Average DC Offset Values for Nine Small Digital Audio Recorders," in *AES 46th International Conference*, Denver, CO, USA, June 2012.

[16] B. Koenig and D. Lacey, "The Average Direct Current Offset Values for Small Digital Audio Recorders in an Acoustically Consistent Environment," *Journal of Forensic Sciences*, vol. 59, no. 4, pp. 960-966, July 2014.

[17] C. Grigoras, D. Rappaport, and J. Smith, "Analytical Framework for Digital Audio Authentication," in *AES 46th International Conference*, Denver, CO, USA, 2012.

[18] C. Grigoras, "Statistical Tools for Multimedia Forensics: Compressoin Effects Analysis," in *39th International Conference of the Audio Engineering Society*, Hillerod, DENMARK, 2010.

[19] S. Hicsonmez and et. al., "Audio codec identification from coded and transcoded audios," *Digital Signal Processing 23*, pp. 1720–1730, 2013.

[20] R. Korycki, "Authenticity examination of lossy compressed digital audio recordings," in *Forum Acusticum*, Krakow, 2014.

[21] D. Gartner, L. Cuccovillo, S. Mann, and P. Aichroth, "A multi-codec audio dataset for codec analysis and tampering detection," in *54th International Conference of the Audio Engineering Society*, London, 2014.

[22] C. Grigoras, "Applications of ENF Analysis in Forensic Authentication of Digital Audio and Video Recordings," *Journal of the Audio Engineering Society*, vol. 57, no. 9, pp. 643-661, 2009.

[23] A. Cooper, "An Automated Approach to the Electric Network Frequency (ENF) Criterion: Theory and Practice," *The International Journal of Speech, Language and the Law*, vol. 16, no. 2, pp. 193-218, 2009.

[24] A. Cooper, "Detecting Butt-Spliced Edits in Forensic Digital Audio Recordings," in *39th International Conference of the Audio Engineering Society*, Hillerod, DENMARK, 2010.

[25] *Standard Practice for Reporting Opinions of Scientific or Technical Experts*, ASTM Standard E620 - 11.

APPENDIX A: SAMPLE CASE

In this sample case the scenario is as follows: Laboratory ‘A’ received a submission of a compact disc (CD). The CD is purported to contain a conversation between two parties. One party contends that the recording is continuous during the entire conversation. The other party contends that the device was paused during recording and it should therefore not be admitted as evidence.

1. Process and document the physical evidence items according to laboratory quality assurance policy.

2. Flatbed scan of submitted CD

3. Hash created for submitted CD

4. Bit-stream images produced and hashed. One placed in archive with the submitted evidence, other used as working copy.

5. Evidence file was identified as Olympus-WS-320050.WMA. Properties reviewed and file hashed.
   MD5: 21ac0bcd0906cb99d9f8a81cc7dda4b8
   SHA256: 7901d8416860de1c4ed3bdea3527ad8b4172d9e4edaae14fc8eea78f2bd00885

6. File structure viewed through hex editor. Determined through header the following pertinent information:
   <u>Structure analysis</u>

File name: Olympus-WS-320050.WMA

```
Offset(h) 00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F
00000000  30 26 B2 75 8E 66 CF 11 A6 D9 00 AA 00 62 CE 6C  0&²uŽfÏ.¦Ù.ª.bÎl
00000010  CE 05 00 00 00 00 00 00 05 00 00 00 01 02 40 A4  Î.............@¤
00000020  D0 D2 07 E3 D2 11 97 F0 00 A0 C9 5E A8 50 12 04  ÐÒ.ãÒ.—ð. É^¨P..
00000030  00 00 00 00 00 00 01 00 10 00 4F 00 4C 00 59 00  ..........O.L.Y.
00000040  4D 00 50 00 55 00 53 00 00 00 01 00 E2 03 02 64  M.P.U.S.....â..d
00000050  73 73 01 00 01 00 02 00 00 00 57 53 33 32 31 4D  ss........WS321M
00000060  20 20 20 20 20 20 20 20 20 20 32 00 00 00 FE FF            2...þÿ
00000070  FF FF FF FF 31 32 31 31 31 33 31 31 32 36 30 37  ÿÿÿÿ121113112607
00000080  31 32 31 31 31 33 31 31 32 36 34 33 30 30 30 30  1211131126430000
00000090  30 34 FF 07 FF FF FF FF FF FF FF FF FF FF FF FF  04ÿ.ÿÿÿÿÿÿÿÿÿÿÿÿ
…
Offset(h) 00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F

00000560  01 00 00 00 02 00 17 00 57 00 69 00 6E 00 64 00  ........W.i.n.d.
00000570  6F 00 77 00 73 00 20 00 4D 00 65 00 64 00 69 00  o.w.s. .M.e.d.i.
00000580  61 00 20 00 41 00 75 00 64 00 69 00 6F 00 20 00  a. .A.u.d.i.o. .
00000590  56 00 38 00 00 00 19 00 20 00 36 00 34 00 20 00  V.8..... .6.4. .
000005A0  6B 00 62 00 70 00 73 00 2C 00 20 00 34 00 34 00  k.b.p.s.,. .4.4.
000005B0  20 00 6B 00 48 00 7A 00 2C 00 20 00 73 00 74 00   .k.H.z.,. .s.t.
000005C0  65 00 72 00 65 00 6F 00 00 00 02 00 61 01 36 26  e.r.e.o.....a.6&
```

Determined through header the following pertinent information:
Type of file: WMA
Recording Device Make: O.L.Y.M.P.U.S
Codec Signature: dss
Device Model: WS321M
Format & Encoder version: Windows Media Audio v.8
Bit rate: 64 kbps
Sampling frequency: 44 KHz
Number of channels: stereo

Note: Not all header information if discussed in this example. If this had been edited, then the hex editor would have shown this. Olympus information would have been stripped and editing software information shown. This is for Olympus WMA files only.

7. Technical characteristics of file documented

Format analysis of Submission

Olympus-WS-320050.WMA Submission	
General	
Format	: Windows Media
File size	: 39.6 KiB
Duration	: 4 s 506 ms
Overall bit rate	: 72.0 kb/s
Maximum Overall bit rate	: 64.0 kb/s
OLYMPUS	: (Binary)
Audio	
ID	: 1
Format	: WMA
Format version	: Version 2
Codec ID	: 161
Codec ID/Info	: Windows Media Audio
Description of the codec	: Windows Media Audio V8 - 64 kbps, 44 kHz, stereo
Duration	: 4 s 506 ms
Bit rate	: 64.0 kb/s
Channel(s)	: 2 channels
Sampling rate	: 44.1 kHz
Bit depth	: 16 bits
Stream size	: 35.2 KiB (89%)

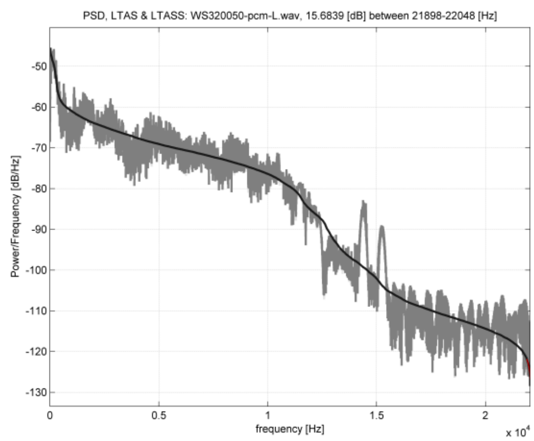
Format analysis shows some of the same information as the hex editor, plus other information decoded from the file metadata.

8. Quantization level (bit-depth) was analyzed indicating bit-depth consistent with an original WMA file made on the submitted recorder (after bit-streaming the memory) or a test recorder of the same make/model (Olympus WS321M).

9. DC offset was analyzed to see if the recording is consistent with a built-in or external microphone, or for indications of editing or processing.

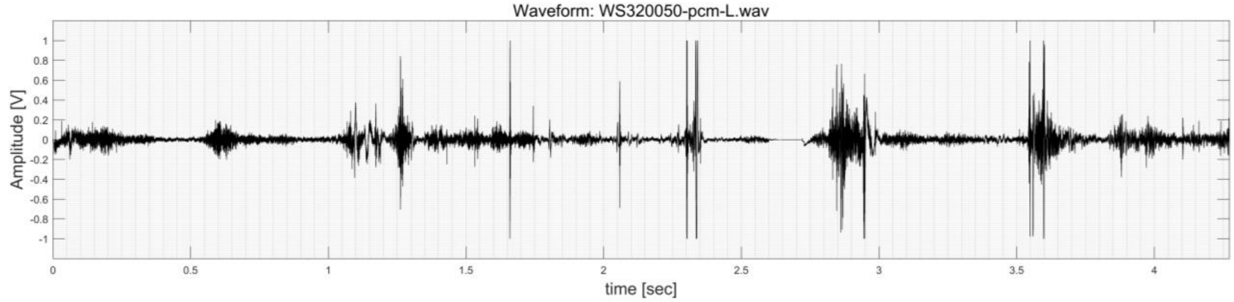
10. Longterm Average Spectral data was produced from an uncompressed conversion of the left channel of the submitted recording. (Power Spectral Density, Long Term Average Spectrum, Long Term Average Sorted Spectrum)

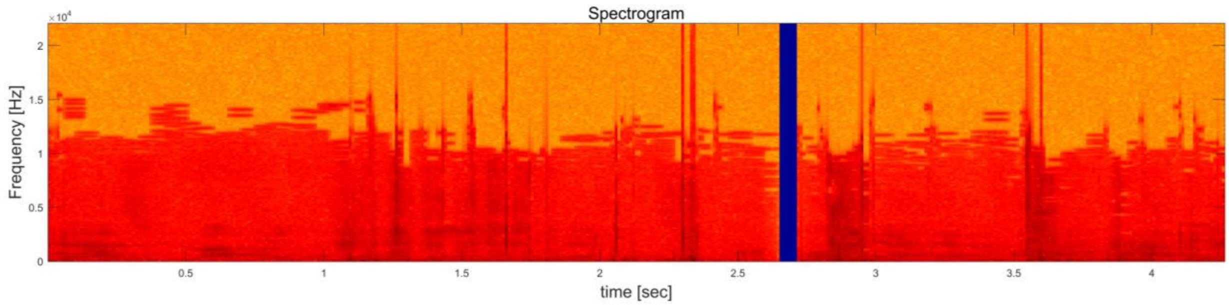
WS-320050.WMA



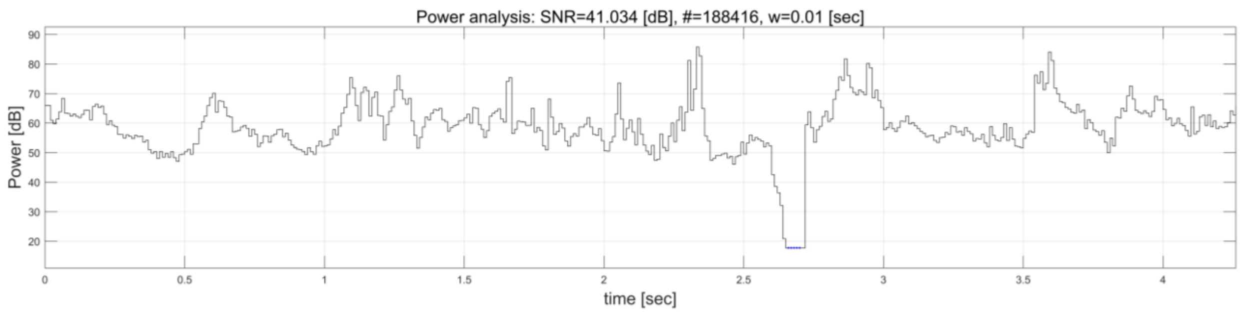
11. No ENF component was detected.

12. Waveform and Spectrogram analysis was performed from an uncompressed conversion of the left channel of the submitted recording, revealing indications of a pause/restart event (left channel).





13. Power analysis of an uncompressed conversion of the left channel of the submitted recording displays pause/restart event.



14. Test recordings were created with a lab-owned Olympus WS-321M digital voice recorder. The following test recordings were produced:

- 40 second continuous recording.

- 30 second recording, pause for 10 seconds, re-started and continued for ten seconds, stopped.

- 30 second recording. Stop. 10 second recording.

- Additionally, a reference file containing a .WMA file that was produced in Adobe Audition by opening a copy of the submitted WMA recording, removing a portion of the recording, and then saving the edited version to a new WMA file..

15. All processes performed on the submitted recording were performed on the test recordings.

Note: Information derived from the 40 second continuous recording and the 30 second recording with stop were consistent with the submission, without the pause/restart event. No notes or graphs from those tests are included in this sample.

Test file With Pause/Restart description

Filename: WS320054.wma
Filesize: 256706 bytes
MD5: b4e3140b20d0729d4f8eb7e3fe788364
SHA256: 27081a101d0a80fa8d992fe53e51ae14f3ac89bda8360436446c36ba5b5e8365

Test file Edited with Adobe Audition description:

Filename: WS320050-AA.wma
Edited in Adobe Audition
Filesize: 44175 bytes
MD5: 45d1c46cf1e806b14fe6fcc68b8219c4
SHA256: 8e8bb8c0f6f69447c134859696b55ead1f251b2cc6330027e941de2c59296809

Structure analysis

Olympus-WS-320050.WMA-Submission	WS320054.WMA-Test with Pause	WS320050-AA.WMA-Test Edited with Adobe Audition
Ofs: 0 -> 0&² Ofs: 3A -> OLYMPUS Ofs: 4F -> dss Ofs: 5A -> WS321M Ofs: 568 -> WindowsMediaAudioV8 Ofs: 59A -> 64kbps Ofs: 5AC -> 44kHz Ofs: 5BC -> stereo --- Time stamp decoding: START REC => 12/11/13 11:26:07 STOP REC => 12/11/13 11:26:43 REC LENGTH => 00:00:04 SESSION LENGTH => 00:00:36 TOTAL PAUSE LENGTH => 00:00:32	Ofs: 0 -> 0&² Ofs: 3A -> OLYMPUS Ofs: 4F -> dss Ofs: 5A -> WS321M Ofs: 568 -> WindowsMediaAudioV8 Ofs: 59A -> 64kbps Ofs: 5AC -> 44kHz Ofs: 5BC -> stereo --- Time stamp decoding: START REC => 12/11/13 12:58:21 STOP REC => 12/11/13 12:59:02 REC LENGTH => 00:00:31 SESSION LENGTH => 00:00:41 TOTAL PAUSE LENGTH => 00:00:10	Ofs: 0 -> 0&² Ofs: 40 -> ToolName Ofs: 56 -> AdobeAudition Ofs: 78 -> WMAFilter Ofs: 96 -> ToolVersion Ofs: CA -> WMFSDKVersion Ofs: 10C -> WMFSDKNeeded Ofs: 142 -> IsVBR Ofs: 253 -> IsVBR Ofs: 26D -> DeviceConformanceTemplate Ofs: 12C1 -> WindowsMediaAudio92 Ofs: 12F5 -> 64kbps Ofs: 1307 -> 44kHz Ofs: 1317 -> stereo Ofs: 1329 -> passCBR

Format analysis

Olympus-WS-320050.WMA-Submission	WS320054.WMA-Test with Pause	WS320050-AA.WMA-Test Edited with Adobe Audition
General Format : Windows Media File size : 39.6 KiB Duration : 4s 506ms Overall bit rate : 72.0 Kbps Maximum Overall bit rate : 64.0 Kbps OLYMPUS : (Binary)	General Format : Windows Media File size : 251 KiB Duration : 31s 142ms Overall bit rate : 65.9 Kbps Maximum Overall bit rate : 64.0 Kbps OLYMPUS : (Binary)	General Format : Windows Media File size : 43.1 KiB Duration : 4s 283ms Overall bit rate mode : Constant Overall bit rate : 82.5 Kbps Maximum Overall bit rate : 64.6 Kbps Encoded date : UTC 2015-06-03 20:59:41.319 Writing application : Adobe® Audition™ WMA Filter 3.0.7283.0
Audio ID : 1 Format : WMA Format version : Version 2 Codec ID : 161 Codec ID/Info : Windows Media Audio Description of the codec : Windows Media Audio V8 - 64 kbps, 44 kHz, stereo Duration : 4s 506ms Bit rate : 64.0 Kbps Channel(s) : 2 channels Sampling rate : 44.1 KHz Bit depth : 16 bits Stream size : 35.2 KiB (89%)	Audio ID : 1 Format : WMA Format version : Version 2 Codec ID : 161 Codec ID/Info : Windows Media Audio Description of the codec : Windows Media Audio V8 - 64 kbps, 44 kHz, stereo Duration : 31s 141ms Bit rate : 64.0 Kbps Channel(s) : 2 channels Sampling rate : 44.1 KHz Bit depth : 16 bits Stream size : 243 KiB (97%)	Audio ID : 1 Format : WMA Format version : Version 2 Codec ID : 161 Codec ID/Info : Windows Media Audio Description of the codec : Windows Media Audio 9.2 - 64 kbps, 44 kHz, stereo 1-pass CBR Duration : 4s 282ms Bit rate mode : Constant Bit rate : 64.0 Kbps Channel(s) : 2 channels Sampling rate : 44.1 KHz Bit depth : 16 bits Stream size : 33.5 KiB (78%) Language : English (US)

File name: WS320054.wma Test with Pause

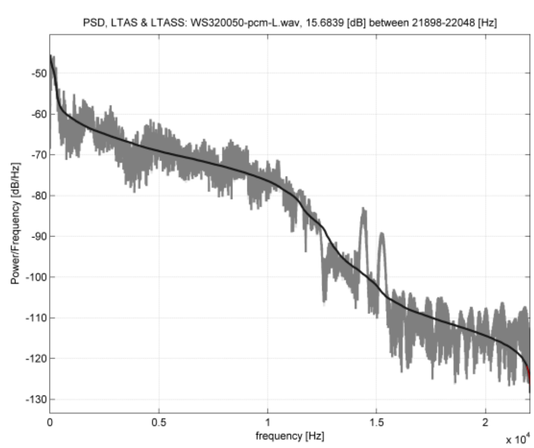
```
Offset(h) 00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F
00000000  30 26 B2 75 8E 66 CF 11 A6 D9 00 AA 00 62 CE 6C  0&²uŽfÏ.¦Ù.ª.bÎl
00000010  CE 05 00 00 00 00 00 00 05 00 00 00 01 02 40 A4  Î.............@¤
00000020  D0 D2 07 E3 D2 11 97 F0 00 A0 C9 5E A8 50 12 04  ÐÒ.ãÒ.—ð. É^¨P..
00000030  00 00 00 00 00 00 01 00 10 00 4F 00 4C 00 59 00  ..........O.L.Y.
00000040  4D 00 50 00 55 00 53 00 00 00 01 00 E2 03 02 64  M.P.U.S.....â..d
00000050  73 73 01 00 01 00 02 00 00 00 57 53 33 32 31 4D  ss........WS321M
00000060  20 20 20 20 20 20 20 20 20 20 36 00 00 00 FE FF            6...þÿ
00000070  FF FF FF FF 31 32 31 31 31 33 31 32 35 38 32 31  ÿÿÿÿ121113125821
00000080  31 32 31 31 31 33 31 32 35 39 30 32 30 30 30 30  1211131259020000
00000090  33 31 FF 07 FF FF FF FF FF FF FF FF FF FF FF FF  31ÿ.ÿÿÿÿÿÿÿÿÿÿÿÿ
```

File name: WS320050-AA.wma Test Edited with Adobe Audition

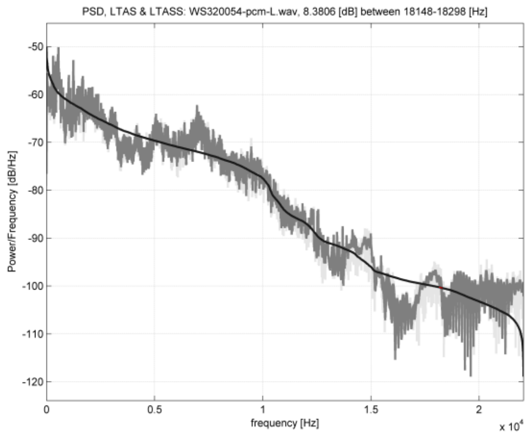
```
Offset(h) 00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F
00000000  30 26 B2 75 8E 66 CF 11 A6 D9 00 AA 00 62 CE 6C  0&²uŽfÏ.¦Ù.ª.bÎl
00000010  D1 13 00 00 00 00 00 00 06 00 00 00 01 02 40 A4  Ñ.............@¤
00000020  D0 D2 07 E3 D2 11 97 F0 00 A0 C9 5E A8 50 38 01  ÐÒ.ãÒ.—ð. É^¨P8.
00000030  00 00 00 00 00 00 05 00 18 00 57 00 4D 00 2F 00  ..........W.M./.
00000040  54 00 6F 00 6F 00 6C 00 4E 00 61 00 6D 00 65 00  T.o.o.l.N.a.m.e.
00000050  00 00 00 00 38 00 41 00 64 00 6F 00 62 00 65 00  ....8.A.d.o.b.e.
00000060  AE 00 20 00 41 00 75 00 64 00 69 00 74 00 69 00  ®. .A.u.d.i.t.i.
00000070  6F 00 6E 00 22 21 20 00 57 00 4D 00 41 00 20 00  o.n."! .W.M.A. .
00000080  46 00 69 00 6C 00 74 00 65 00 72 00 00 00 1E 00  F.i.l.t.e.r.....
00000090  57 00 4D 00 2F 00 54 00 6F 00 6F 00 6C 00 56 00  W.M./.T.o.o.l.V.
000000A0  65 00 72 00 73 00 69 00 6F 00 6E 00 00 00 00 00  e.r.s.i.o.n.....
000000B0  16 00 33 00 2E 00 30 00 2E 00 37 00 32 00 38 00  ..3...0...7.2.8.
000000C0  33 00 2E 00 30 00 00 00 1C 00 57 00 4D 00 46 00  3...0.....W.M.F.
000000D0  53 00 44 00 4B 00 56 00 65 00 72 00 73 00 69 00  S.D.K.V.e.r.s.i.
000000E0  6F 00 6E 00 00 00 00 00 20 00 31 00 32 00 2E 00  o.n..... .1.2...
000000F0  30 00 2E 00 37 00 36 00 30 00 31 00 2E 00 31 00  0...7.6.0.1...1.
00000100  37 00 35 00 31 00 34 00 00 00 1A 00 57 00 4D 00  7.5.1.4.....W.M.
00000110  46 00 53 00 44 00 4B 00 4E 00 65 00 65 00 64 00  F.S.D.K.N.e.e.d.
00000120  65 00 64 00 00 00 00 00 16 00 30 00 2E 00 30 00  e.d.......0...0.
00000130  2E 00 30 00 2E 00 30 00 30 00 30 00 30 00 00 00  ..0...0.0.0.0...
00000140  0C 00 49 00 73 00 56 00 42 00 52 00 00 00 02 00  ..I.s.V.B.R.....
```

Longterm Average Spectral data was produced. (Power Spectral Density, Long Term Average Spectrum, Long Term Average Sorted Spectrum)

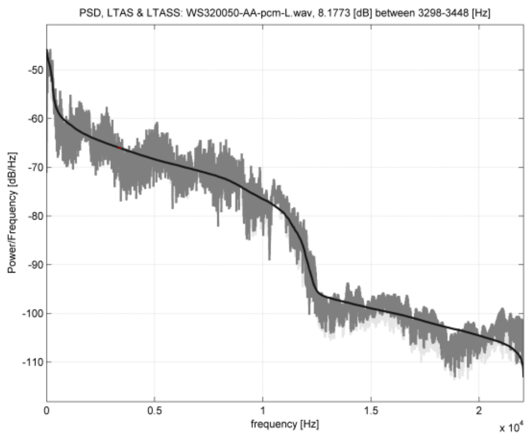
Olympus-WS-320050.WMA-Submission



WS320054.WMA-Test with Pause



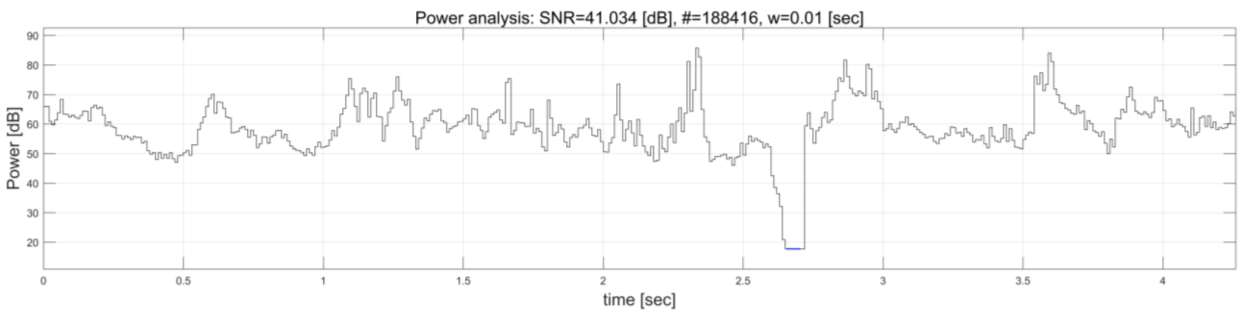
WS320050-AA.WMA-Test Edited with Adobe Audition



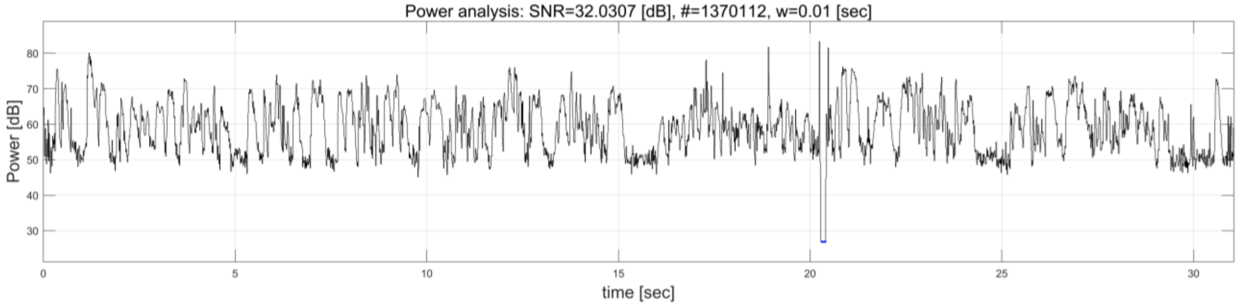
Waveform and Spectrogram Analysis of test recordings support Pause Event.

Power analysis confirms Pause Event.

WS320050.WMA-Submission



WS320054.WMA-Test with Pause



16. Report was drafted appropriate for the jurisdiction. The conclusions were summarized with these statements and the following table:
    a. Submitted file Olympus-WS-320050.WMA is consistent with an original recording made by an Olympus WS-321M. A Pause/re-start event is present at approximately 2.7 seconds.

Sample Case Authenticity Analysis Results – 'Olympus-WS-320050.WMA'		
Category	Analysis	Result
Global Analysis	Structure	*consistent*
	Format	*consistent*
	Quantization Level (Bit Depth)	*consistent*
	DC Offset	*consistent*
	Long Term Average Spectrum	*consistent*
	Sorted Spectrum	*consistent*
	Electric Network Frequency (ENF) Analysis	*N/A*
Local Analysis (event at 2.7 seconds against test pause/restart event)	Waveform Analysis	*consistent*
	Spectrogram Analysis	*consistent*
	Signal's Power	*consistent*

Best Practices for Digital Audio Authentication

History

Rev	Issue Date	Section	History
1.01	2011-09-15	All	Document initiated.
1.32	2015-06-01	All	SWGDE meeting development.
1.33	2015-06-04	All	Approved by SWGDE for release as a Draft for Public Comment.
1.34	2015-06-30	All	Formatting and tech edit performed for release as a Draft for Public Comment; published on SWGDE website.
1.42	2016-01-14	All	Updated based on public comments on the draft.
1.42	2016-01-14	All	Approved by SWGDE for re-release as a Draft for Public Comment.
1.43	2016-02-08	All	Formatting and technical edit performed for release as a Draft for Public Comment.
1.46	2016-06-09	All	Updated in Audio Committee to incorporate public comments, references, and discussions. SWGDE voted to publish as an Approved document.
1.47	2016-06-23	--	Formatted and posted as an Approved document, Version 1.
1.48	2016-09-15	3.3, 4.3, 4.4	Language edits for clarity, consistency. SWGDE voted to publish as an Approved document.
1.49	2016-10-06	4.4.2	Provided for formatting and posting as Approved Version 1.1.
Ver. 1.1	2016-10-08	--	Formatted to publish as Approved version 1.1.
1.50	2017-01-12	2.	Editorial change to “original evidence” definition – subject/verb agreement. Changed: “Not all original recordings are recoverable from their device.” To: “Not every original recording is recoverable on its storage device.” SWGDE voted to publish as an Approved document, Version 1.2.
Ver. 1.2	2017-02-21	--	Formatted and published as Approved version 1.2.
Ver. 1.3	2018-09-20	6	Added references: [11], [12], and [13]. Re-published as Approved version 1.3.