**** **ISO/IEC JTC 1/SC 29/WG 6 N337**

**ISO/IEC JTC 1/SC 29/WG 6**

**MPEG Audio Coding Convenorship: DE**

**Document type:** Output Document

**Title:** Call for proposals on Audio Coding for Machines

**Status:** Approved

**Date of document:** 2025-04-04

**Source:** ISO/IEC JTC 1/SC 29/WG 6

**Expected action:** None

**Action due date:** None

**No. of pages:** 12 (without cover page)

**Email of Convenor:** Thomas.Sporer@idmt.fraunhofer.de

**Committee URL:** <https://isotc.iso.org/livelink/livelink/open/jtc1sc29wg6>

**INTERNATIONAL ORGANISATION FOR STANDARDISATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC 1/SC 29/WG 6**

**MPEG AUDIO CODING**

**ISO/IEC JTC 1/SC 29/WG 6 N337**

**April 2025, Virtual**

|  |  |
| --- | --- |
| **Title** | **Call for proposals on Audio Coding for Machines** |
| **Source** | **WG 6, MPEG Audio Coding** |
| **Status** | **Approved** |
| **Serial Number** | **25009** |

This document explains all necessary information for the Call for Proposals on Audio Coding for Machines including links to supporting documents.

**Table of Contents**

[1 Introduction 4](#_Toc194687773)

[1.1 System Overview 4](#_Toc194687774)

[2 Who may participate 5](#_Toc194687775)

[3 Code of conduct and rules of engagement 6](#_Toc194687776)

[4 Source code and IPR 6](#_Toc194687777)

[5 Testing Fee 6](#_Toc194687778)

[6 Definitions 6](#_Toc194687779)

[7 Documents of CfP package 6](#_Toc194687780)

[8 Submission Process 6](#_Toc194687781)

[8.1 CfP Timeline table 6](#_Toc194687782)

[8.2 Envisioned Timeline for the ACoM Standard 7](#_Toc194687783)

[8.3 Register your participation 7](#_Toc194687784)

[8.4 Mandatory Equipment, Software and Data Components 7](#_Toc194687785)

[8.5 Access to Test Material 8](#_Toc194687786)

[8.6 Conduct Objective Evaluations 8](#_Toc194687787)

[8.7 Submit Proponent Documentation 8](#_Toc194687788)

[8.8 Evaluate CfP Submissions and Select Technology 8](#_Toc194687789)

[8.9 Submit WD Specification and RM Source Code 9](#_Toc194687790)

[9 Call Administrator 9](#_Toc194687791)

[10 References 9](#_Toc194687792)

[11 Copyright Header for ACoM Reference Software 9](#_Toc194687793)

[Appendix A: Evaluation of RESULTS 11](#_Toc194687794)

[1 Requirements 11](#_Toc194687795)

[1.1 Mandatory Requirements 11](#_Toc194687796)

[1.2 Additional Requirements 11](#_Toc194687797)

[2 Anchor 11](#_Toc194687798)

[3 Evaluation in the CfP 12](#_Toc194687799)

[4 Analysis of data and selection of technology 12](#_Toc194687800)

[4.1 Overview 12](#_Toc194687801)

[4.2 Figure of merit 12](#_Toc194687802)

[4.2.1 Fulfilling requirements 12](#_Toc194687803)

[4.2.2 Comparison of performance 12](#_Toc194687804)

[4.2.3 Merging of winning technologies 13](#_Toc194687805)

# Introduction

Traditional coding methods aim for the best audio under certain bit-rate constraint for human consumption. However, with the rise of machine learning applications, along with the abundance of sensors, many intelligent platforms have been implemented with massive data requirements including scenarios such as connected vehicles, audio surveillance, machine diagnostics, health monitoring, artistic creation, and smart city. In many of these applications the spatial distribution of audio contains important information. Medical data, like those coming from EEG and ECG measurements, while not inherently audio, often have a similar structure and similar requirements as audio data. Spatial audio and medical data will be referenced as multi-dimensional streams.

The sheer quantity of data being produced constantly leads previous methods with a human in the pipeline to be inefficient, and unrealistic in terms of latency and scale. There are additional concerns in transmission and archive systems which require a more compact data representation and low latency solution. This led to the introduction of Audio Coding for Machines.

In some cases, machines will communicate amongst themselves to perform tasks without a human in the mix, while in others there will be a need for additional human consumption of the specific decompressed stream. This specific scenario is possible in surveillance use cases, where a human “supervisor” may occasionally search for a specific sound, or scene in the audio. In other cases, the corresponding bitstream may be used for both human and machine consumption. In the case of cars, the features may be used to overcome the good sound insulation of modern cars (“detection of ambulance sirens”) and for monitoring failure of components (including predicted maintenance).

Any use cases in which audio features need to be transmitted for additional processing which may potentially be used for machine or human end users could benefit from a standard in the coded features (shared backbone). Interoperability is crucial where different manufacturers and platforms need communication to achieve a common goal.

Additionally, the feature stream must be efficient for both transmission and archive concerns for both latency and space. A standard for the compressed coding of this feature stream will establish an efficient protocol for machines to communicate.

MPEG-ACoM aims to define a bitstream and data format for compressing audio, multi-dimensional streams, or features extracted from such signals that is efficient in terms of bitrate/size and can be used by a network of machines after decompression to perform multiple tasks without significantly degrading task performance. The decoded audio, multi-dimensional streams or features can be used for machine consumption or hybrid machine and human consumption. In addition to the essence the format must also contain metadata describing how the data audio or multi-dimensional stream was captured. The first phase should be application agnostic: Data is encoded near-lossless enabling the training of feature extraction schemes. The result from this phase is already useful for industry simplifying the exchange of data using this standardized format. In a second phase feature extraction schemes are added. These features will be optimized for different applications.

## System Overview

The generic system architecture contains a pair of ACoM encoder and decoder. The input of the ACoM system could be metadata describing the input and either of

* audio signals (one or multi-dimensional)
* multi-dimensional streams (e.g. medical data)
* extracted features (phase 2 only).

In case of a feature stream, the type and format of feature should be specified, features may take different forms depending on the application. Feature extraction and coding is not in phase 1 but in phase 2.

The decompressed bitstream of audio and/or multi-dimensional streams and/or features may then be used for post-processing tasks, which may include machine consumption tasks or hybrid machine and human consumption tasks. The encoder can be optimized for either a single task or multiple, and the size of the compressed stream should compare favorably to traditional coding techniques on the unprocessed audio or medical data.

The MPEG activity on Audio Coding for Machines (ACoM) aims to standardize a bitstream format generated by compressing a previously extracted feature stream or data stream. Compression scheme of data stream including the audio and metadata at each phase are followed:

|  |  |  |
| --- | --- | --- |
| Phase | Data Type | Compression Scheme |
| Phase 1 | Audio/Signal | Lossless compression |
| Metadata | Lossy/lossless compression |
| Phase 2 | Audio/Signal | Lossy/lossless compression |
| Metadata | Lossy/lossless compression |

A diagram of a rectangular object with arrows

AI-generated content may be incorrect.

**Figure 1:** Pipeline for ACoM: Phase 1 using lossless compression. In Phase 2, compression is expected to use (lossy) feature compression.

A diagram of a computer

AI-generated content may be incorrect.

**Figure 2**: An example of potential ACoM architecture (both phase 1 and phase 2).

Fig 2 shows an example of potential ACoM architecture. The ACoM codec could be an audio codec or a feature codec, or both. In case of a feature codec, the ACoM feature encoding is consisting of feature extraction, feature conversion and feature coding. In case of an audio codec, an adaptive pre-processing may be included in ACoM encoder, and an adaptive post-processing may be included in ACoM decoder. Metadata may be fed into both audio and metadata encoding modules. There may be an interface to an external neural network (NN) for the feature extraction and the task specific networks.

This document contains detailed information about the setup of this call, general rules for conduct, the planned timeline for both this call and the development of a standard, detailed information on what needs to be submitted, and the next steps following the evaluation of responses to the call.

# Who may participate

Proponents that respond to this call may include any persons whether they are or are not accredited delegates of ISO/IEC JTC1/SC29/WG 6. However, all proponents are required to attend the meetings at which their respective proposals are evaluated. The meeting during which proposals are evaluated is identified with \* in Table 1. A one-time invitation may be extended to proponents to participate in the evaluation process if the proponent is not an accredited delegate of ISO/IEC JTC1/SC29/WG 6. If the proponent’s technology is accepted into the Working Draft of the Standard, then the proponents are required to participate in meetings identified with a † in Table 1 and Table 2. In such a case where the technology is accepted from a proponent who is not an accredited delegate of ISO/IEC JTC1/SC29/WG 6, the proponent is expected to initiate the process to join their National Body committees in order to become accredited to participate in subsequent meetings of WG 6. Information for how to join National Body committees and to become an accredited delegate for ISO/IEC JTC1/SC29/WG 6 is available at [How to Get Involved](https://www.iso.org/get-involved.html).

# Code of conduct and rules of engagement

All participants shall be required to familiarize themselves with relevant [ISO Policies and Procedures](https://www.iso.org/resources.html), including in particular [ISO Code of Conduct](https://www.iso.org/publication/PUB100397.html), [ISO Declaration for Participants in ISO Activities](https://www.iso.org/declaration-for-participants-in-iso-activities.html), [ISO Privacy and Copyright](https://www.iso.org/privacy-and-copyright.html) policy, and [ISO Policy on Communication of Committee Work](https://www.iso.org/publication/PUB100382.html), and to consent to be bound by these policies.

# Source code and IPR

By responding to a CfP, the proponent affirms that he or she is willing to make source code available for use as the starting point for collaborative standardization.

It is the responsibility of the proponent to obtain any necessary internal approvals in a timely manner, otherwise more readily available source code may be selected.

Furthermore, proponents are advised that this Call is being made subject to the common patent policy of ITU-T/ITU-R/ISO/IEC (refer to [www.itu.int/ITU-T/dbase/patent/patent-policy.html](http://www.itu.int/ITU-T/dbase/patent/patent-policy.html) or Appendix I of [ISO/IEC Directives Part 1](http://isotc.iso.org/livelink/livelink?func=ll&objId=4230455&objAction=browse&sort=subtype)).

# Testing Fee

Participating in this Call for Proposals is not associated with any fees.

# Definitions

The definitions for terms associated with this Call for Proposals are found in [1].

Furthermore, this Call for Proposals uses the terms Working Draft (WD), Committee Draft (CD), Draft International Standard (DIS), Final Draft International Standard (FDIS) and International Standard (IS) according to the [ISO Stages and Resources for Standard Development](https://www.iso.org/stages-and-resources-for-standards-development.html).

# Documents of CfP package

The CfP package consists of the following documents:

* WG6 N337 Call for proposals on audio coding for machines (this document)

This document contains details about the submission process and the guidelines to follow. Furthermore, it describes who may participate and what the rules for participation are. In Annex of this document, figure of merit and how to obtain test sequences are given.

* WG6 N339 List of content available for CfP of audio coding for machines
* WG6 N340 Metadata format for audio coding for machines
* WG6 N341 Rules for core experiments for audio coding for machines

Supplementary information can be found in

* WG2 N453 Use cases and requirements on audio coding for machines

(on <https://mpeg.org> )

# Submission Process

## CfP Timeline table

Each entry in the table is described in a section below. WG 6 is the SC 29 working group WG 6 MPEG Audio Coding. Unless stated otherwise, deadlines refer to a specific day at 23:59 UTC.

Table 1. CfP Timeline (\* indicates attendance at the meeting is required. † indicates attendance is required if technology is selected to be included in the Working Draft).

|  |  |  |  |
| --- | --- | --- | --- |
| **Meeting** | **Date** | **Who** | **Action** |
| 19 | April 2025 | WG2/WG6 | Issue Call for Proposals package |
| 19 | April 2025 | WG6 | Verified version of test material is available |
|  | Aug 1, 2025 | Proponent | Register proposals |
|  | Sep 1, 2025 | Proponent | Upload bitstream files, results and decoders |
|  | Sep 8, 2025 | Volunteers and Proponents | Begin cross-checking others' results. |
|  | Sep 22, 2025 | Volunteers and Proponents | Complete cross-checking others results |
|  | Sep 26, 2025 | Volunteers and Proponents | Submit results of cross-checking as a contribution to the 21st WG 6 meeting |
|  | Sep 26, 2025 | Proponent | Submit proponent documentation as a contribution to the 21st WG 6 meeting |
| 21\* | Oct. 2025  (10/7-10/11) | WG 6 | Evaluate Call for Proposals submissions and select technology. Respondents must be present at the meeting and present the proposals for information sharing.  The source code of candidate technologies for the reference model will be published. |
|  | Nov. 7, 2025 | Software coordinators | Source code of the selected technologies published on Git |
| 22† | Jan. 2026 | AHG | Submit reference model (RM) for selected technologies as input contribution. |
| 22† | Jan. 2026 | AHG | Submit Working Draft (WD) for selected technologies as input contribution. |

## Envisioned Timeline for the ACoM Standard

It is envisioned that the timetable for the progress of the ACoM standard will be as follows. Note that not all exact dates for future meetings are set and AhG meetings might be in advance of the WG 6 meetings.

Table 2. Standardization Timeline (\* indicates attendance at the meeting is required. † indicates attendance is required if technology is selected to be included in the Working Draft.)

|  |  |  |
| --- | --- | --- |
| **WG 6 Mtg** | **Date** | **Action** |
| 19 | March 31-April 4, 2025 | CfP |
| 21\* | October 7 – 11, 2025 | Evaluation |
| 22† | January 19-23, 2026 | RM |
| 22† | January 19-23, 2026 | WD |
| 23† | April 27 - May 1, 2026 |  |
| 24† | July 13 – 17, 2026 |  |
| 25† | October 19-23, 2026 | CD |
| 26† | January 18-22, 2027 | DIS |
| 28† | July, 2027 | Verification Test |
| 28† | July, 2027 | FDIS |
| 30 | January, 2028 | IS |

## Register your participation

Proponent must register on or before the date shown in the CfP timeline table above, an intention to participate in the CfP. Registering an intent is not binding and registered parties are not required to submit proposals. However, parties that do not register will not be able to submit proposals. Register by sending an email to Thomas Sporer (Convenor, WG 6 MPEG Audio Coding, Thomas.sporer@idmt.fhg.de). Email should indicate

* Company name
* Contact name and contact email address
* Every proposal should handle for every use cases, but proponents are invited to notice about their own weighting of use cases (for later use).

## Mandatory Equipment, Software and Data Components

* The decoder shall be assessed on Linux running on x86 platform.
* The baseline software shall be installed on Linux running on x86 platform. For details on software releases see Annex A “Figure of merit”

## Access to Test Material

* Test materials is at MPEG git (https://git.mpeg.expert/MPEG/audio/ACoM/cfp ).
* For git access, first you must be registered, then ask permission by email to WG6 Convenor ( Thomas Sporer, Thomas.sporer@idmt.fraunhofer.de ).
* Directory for uploading test items (decoder and bitstreams) will be prepared and named as proposals/companyName.

## Conduct Objective Evaluations

Proponents must provide the objective measurement of the following parameters. Details on how to measure the following metrics in Annex A.

The main evaluation metrics are:

* Bitrate reduction
* Checking of lossless audio
* Checking of metadata coded within the allowed tolerance

The following metrics shall be reported to provide additional information about the proposal and may be used in selecting the technology for the reference model (use later for CE process, WG6 N341):

* decoder runtimes
* decoder size
* encoder runtimes
* encoder size

After registering [see section 8.3], a template for the submission of this information will be provided by WG6 Convenor. The proponent shall include a summary of the objective results in the contribution document describing the proposed technology. This shall include a description about the hardware and software environment where the objective evaluation is performed.

## Submit Proponent Documentation

Proponents submit as a contribution to the MPEG meeting indicated in the CfP Time Line table, above:

* A written description of the technology having sufficient detail to permit technical discussions.
* Objective test results, as indicated in template mentioned in section 8.6.
* Description of how requirements are met.

Proponents that are WG 6 members shall register these documents as contributions to the WG 6 meeting and send title and author information to ACoM email list, prior to the close of contribution upload deadline. All proponents are encouraged to become WG 6 members. However, proponents that are not WG 6 members shall email the documents to Convenor of WG 6 two weeks prior to the WG 6 meeting at which proponent documentation materials are due, so that he can register and upload them as contributions. The documents should be written in Microsoft Word. The Convenor of WG 6 will extend a one-time-only invitation to the WG 6 meeting so that a non-member proponent can present their contributions and participate in the selection process.

After the submission of bitstreams and decoders, each proponent shall cross-check the results reported by other proponents. Every proposal must be cross-checked by other experts of other company. Non-proponent participants might assist with cross-checking. A report on the results of this cross-checking activity shall be made available by the cross-checkers by the date listed in the CfP Timeline table above. This report shall be uploaded as an input contribution to the WG 6 meeting. The Test Administrator assigns and communicates the cross-checking tasks to cross-checkers directly. A cross-checker shall not disclose information about the IDs of their cross-checking assignments or information about the methods in their cross-checking assignments until the cross-checking contribution to the MPEG has been submitted.

The cross-check for a proposal is considered successful if the file sizes of the proposal are matched exactly, and the error of metadata coding within the tolerance and coded audio essence is lossless.

Non-proponents that are interested in participating in the cross-checking effort are welcome to contact the Test Administrator and indicate the number of proposals they would like to cross-check.

## Evaluate CfP Submissions and Select Technology

At the WG 2/ WG 6 meeting indicated in the CfP Timeline table above, submissions will be evaluated by the WG 6 experts. It is strongly urged that proponents have experts familiar with the proposed technology attend in order to allow discussions on details of the proposals. It is envisioned that at least one submission will be selected as technology for the Working Draft of the ACoM standard. Submissions shall be evaluated considering all submitted information.

The metrics for evaluating the submissions can be found in Annex A.

Proposals do not have to fulfill all optional requirements. Requirements that are not fulfilled by the selected technology will be addressed in the Core Experiment (CE) process, which may include CEs using other submitted technologies in order to address additional optional requirements.

If by the assessment of the WG 6 experts, there is no single best proposal, then WG6 will draft a workplan on how to merge the best-performing technologies into a single unified technology.

## Submit WD Specification and RM Source Code

At the WG 6 meeting indicated in the timetable above, the AHG shall submit an initial version for the Working Draft (WD) of the specification as an input contribution. Based on this contribution, the group will collaborate to create a WD.

The WD must include a normative specification of the ACoM decoding process, signaling aspects and bitstream syntax.

The source code RM0 of the selected technologies shall be published in Git within 4 weeks after the WG6 meeting where the selected technology is announced. If the proponent(s) fail to publish the source code on time, new technologies will be selected, and a new work plan will be drafted by the WG6 experts. Software coordinators shall be selected from proponent(s) of the selected technology and/or interested experts. If necessary, the merge activity shall be carried out as an open process (i.e., visible to all members of WG6) on the Git repository. The merging of the code shall be completed before the date indicated in the timeline and a contribution by the software coordinator shall describe the merging process, details of how to access the code, and how to jointly develop the code during the standardization process.

The RM shall be cross-checked by at least 2 independent WG6 member organizations. If a single proposal was chosen, the cross-checking shall verify that

* the bitstreams generated by the encoder in the RM match the bitstreams in the original CfP submission
* the reconstructed sound/audio generated by the decoder in the RM match the reconstructed sound/audio generated by the compiled decoder in the original CfP submission

If the RM is not based on a single proposal, the cross-checking shall produce a set of baseline results for the collaborative development of the ACoM standard.

In addition, the contribution containing the WD shall include a description of the encoding and decoding algorithms. To match this, the RM must include source code that implements the described encoding and decoding algorithms. Note that this does not need to be the exact algorithm used in the proponent’s submission to the CfP but rather the merged software. Subsequent Core Experiment work for the collaborative development of ACoM standard shall use the RM as the “baseline” system in CE performance comparisons*,* against which CE technology is evaluated.

# Call Administrator

For any questions related to this Call for Proposals or associated evaluation procedures please contact the Test Administrator:

Dr. Thomas Sporer, Convenor, WG 6 MPEG Audio Coding

Fraunhofer-IDMT

Ehrenbergstraße 31,

98693 Ilmenau, Germany

Phone: +49 3677 4670

email: Thomas.sporer@idmt.fraunhofer.de

# References

1. WG2 N453 use cases and requirements for audio coding for machines

# Copyright Header for ACoM Reference Software

All ACoM Reference Software files shall contain the following header:

/\* The copyright in this software is being made available under the BSD

\* License, included below. This software may be subject to other third party

\* and contributor rights, including patent rights, and no such rights are

\* granted under this license.

\*

\* Copyright (c) 202x, ISO/IEC

\* All rights reserved.

\*

\* Redistribution and use in source and binary forms, with or without

\* modification, are permitted provided that the following conditions are met:

\*

\* \* Redistributions of source code must retain the above copyright notice,

\* this list of conditions and the following disclaimer.

\* \* Redistributions in binary form must reproduce the above copyright notice,

\* this list of conditions and the following disclaimer in the documentation

\* and/or other materials provided with the distribution.

\* \* Neither the name of the ISO/IEC nor the names of its contributors may

\* be used to endorse or promote products derived from this software without

\* specific prior written permission.

\*

\* THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS"

\* AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE

\* IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE

\* ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT HOLDER OR CONTRIBUTORS

\* BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR

\* CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF

\* SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS

\* INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN

\* CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE)

\* ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF

\* THE POSSIBILITY OF SUCH DAMAGE.

\*/

Appendix A: Evaluation of RESULTS

# Requirements

## Mandatory Requirements

* The bitstream must represent the audio essence, the metadata, and the license file.
* Decoder must generate output files with all metadata fields present in the input file and the license file. It is not necessary to encode and decode the tolerance fields.
* The difference between the input audio file and the output file excluding blanks at begin and end of file must be zero.
* The output metadata file must preserve information of the input metadata file. If in the input metadata file tolerances for pose are given the difference between input and output must be within the tolerance defined in the input file.

## Additional Requirements

There are additional requirements which will be not checked in the call for proposals. They will be considered in the CE procedure. These might be:

* **Length of access unit**

Target value: as short as possible

Note: for streaming application this is related to the total system delay

* **Computational complexity**:

Target: as short as possible

Note: for long time monitoring this value is related to batterie live-time

* **Memory consumption of encoder and decoder**:

Target: as low as possible

# Anchor

The evaluation of performance is based on the relative improvement of each proposal compared to a baseline codec. The baseline codec consists of a codec to lossless encode the audio essence and a lossless compression scheme to bundle the encoded audio essence, the metadata file and the rights file in one file. The following codecs are used:

* For the first step the audio essence mpeg-4 SLS is used.  
  The code can be found in the output document N12794 "ISO/IEC 14496-5:2001/AMD 10:2007/Cor 6, SLS Corrections". In addition, the modules *libisomedia* and *afsp* are needed.

afsp can be found at <https://www.mmsp.ece.mcgill.ca/Documents/Downloads/AFsp/index.html>.

* For the second step info-zip 3.0 is used (<https://infozip.sourceforge.net/>). The compression mode is “*deflate*”, the “compression level is “-6”.
* The length of the resulting file is measured via the command *ls -b*

An example for the toolchain for encoding the files “item1.xml”, “item1.wav”, and “LICENSE.txt” is given here:

**sls\_enc -ncore -o ./item1.mp4 ./item1.wav**

**zip -6 -Z deflate item1.zip item1.xml item1.mp4 LICENSE.txt**

**ls -b item1.zip**

To check correctness of the baseline use the following toolchain:

**mkdir check**

**cd check  
 unzip ../item1.zip**

**sls\_dec -D -w ./item1.mp4 ./item1-out.wav**

**CompAudio item1-out.wav ../item1.wav**

Note: CompAudio is part of the afsp library. This program compares audio files of different formats and is able to ignore leading and trailing zeros.

# Evaluation in the CfP

* Every proposal must include a decoder to decode and generate the files equivalent to raw input (audio files, metadata files, and license files). The decoder must run on Linux on x86 platform. The decoder might consist of different parts which are linked together with scripting language.
* Every proposal must be accompanied with a table expressing the following measurements for each item:
  + File size of the proposed compressed bitstream in Bytes
  + Encoding time for the proposal
  + Decoding time for the proposal

To normalize these values every proposal must also include measurements of the baseline done on the same hardware:

* + File size of the baseline bitstream in Bytes
  + Encoding time for the baseline
  + Decoding time for the baseline

Memory usage of encoder and decoder

* + Proponents shall estimate memory usage and describe the method used.
  + The memory usage shall be estimated and reported by the proponent for each item. They shall express the highest memory load during running the encoder or decoder.

# Analysis of data and selection of technology

# Overview

The selection of a proponent technology as RM0 shall be made by considering objective performance information.

# Figure of merit

The figure of merit consists of two steps: fulfilling requirements and comparison of performance:

# Fulfilling requirements

* For each recording the mandatory requirements are checked: lossless for the audio essence and metadata encoded within the tolerances.
* If mandatory requirements in any of the items are not fulfilled the whole proposal is rejected from further evaluation.

# Comparison of performance

For the Immersive Audio CfP evaluation, the achieved **compression ratio** is taken as the Figure of Merit (FoM).

For each item the ratio between the size of the compressed bitstream and the size of the baseline is calculated.

The FoM shall be the **grand mean over all test items and all use cases**. Additionally, the grand mean is associated with a 95% confidence interval on the mean. The proposal with the highest FoM will be designated as the “**overall *winner***” only if it is significantly better than the other proposals.

Significantly better is defined as pairwise comparison as follows: The means of first proposal is not within the confidence intervals of the other proposal and vie-versa.

If there is a number of proposals not significantly best the worst-case performance is taken into account: Among the proposals with overlapping intervals the proposal with the best worst-case performance will be selected as the winner.

All these calculations are assuming 32bit floating point in IEEE format.

In addition, a similar procedure is performed for each use case separately: If in a use case the number of test items is below 10 the standard deviation of the overall test (all use cases) is used instead of the standard deviation of the items of this use case.   
For a use case a “**use case winner**” is only selected if it is significantly better than the “overall winner”.

# Merging of winning technologies

The “overall winner” over all use cases form RM0 base.

If there are “***use case winners***” (if any) proponents shall work towards a merging into “*winner*” solution.