Moving Picture, Audio and Data Coding
by Artificial Intelligence
www.mpai.community

MPAI Technical Specification

Context-based Audio Enhancement
MPAI-CAE – Use Cases

V2.2

WARNING

Use of the technologies described in this Technical Specification may infringe patents, copyrights or intellectual property rights of MPAI Members or non-members.

MPAI and its Members accept no responsibility whatsoever for damages or liability, direct or consequential, which may result from the use of this Technical Specification.

Readers are invited to review Notices and Disclaimers.

Technical Specification: Context-based Audio Enhancement (MPAI-CAE) – Use Cases V2.2

Contents

1 Foreword

The international, unaffiliated, non-profit ***Moving Picture, Audio, and Data Coding by Artificial Intelligence (MPAI)*** organisation was established in September 2020 in the context of:

1. **Increasing** use of Artificial Intelligence (AI) technologies applied to a broad range of domains affecting millions of people
2. **Marginal** reliance on standards in the development of those AI applications
3. **Unprecedented** impact exerted by standards on the digital media industry affecting billions of people

believing that AI-based data coding standards will have a similar positive impact on the Information and Communication Technology industry.

The design principles of the MPAI organisation as established by the MPAI Statutes are the development of AI-based Data Coding standards in pursuit of the following policies:

1. *Publish* upfront clear Intellectual Property Rights licensing frameworks.
2. *Adhere* to a rigorous standard development process.
3. *Be friendly* to the AI context but, to the extent possible, remain agnostic to the technology thus allowing developers freedom in the selection of the more appropriate – AI or Data Processing – technologies for their needs.
4. *Be attractive* to different industries, end users, and regulators.
5. *Address* five standardisation areas:

1. Data Type, a particular type of Data, e.g., Audio, Visual, Object, Scenes, and Descriptors with as clear semantics as possible.
2. Qualifier, specialised Metadata conveying information on Sub-Types, Formats, and Attributes of a Data Type.
3. AI Module (AIM), processing elements with identified functions and input/output Data Types.
4. AI Workflow (AIW), MPAI-specified configurations of AIMs with identified functions and input/output Data Types.
5. AI Framework (AIF), an environment enabling dynamic configuration, initialisation, execution, and control of AIWs.

6. *Provide* appropriate Governance of the ecosystem created by MPAI Technical Specifications enabling users to:
    1. Operate Reference Software Implementations of MPAI Technical Specifications provided together with Reference Software Specifications
    2. Test the conformance of an implementation with a Technical Specification using the Conformance Testing Specification.
    3. Assess the performance of an implementation of a Technical Specification using the Performance Assessment Specification.
    4. Get conforming implementations possibly with a performance assessment report from a trusted source through the MPAI Store.

Today, the MPAI organisation rests on four solid pillars:

1. The MPAI Patent Policy specifies the MPAI standard development process and the Framework Licence development guidelines.
2. *Technical Specification: Artificial Intelligence Framework (MPAI-AIF)* specifies an environment enabling initialisation, dynamic configuration, and control of AIWs in the standard AI Framework environment depicted in Figure 1. An AI Framework can execute AI applications called AI Workflows (AIW). An AIW includes interconnected AI Modules (AIM). MPAI-AIF supports small- and large-scale high-performance components and promotes solutions with improved explainability.

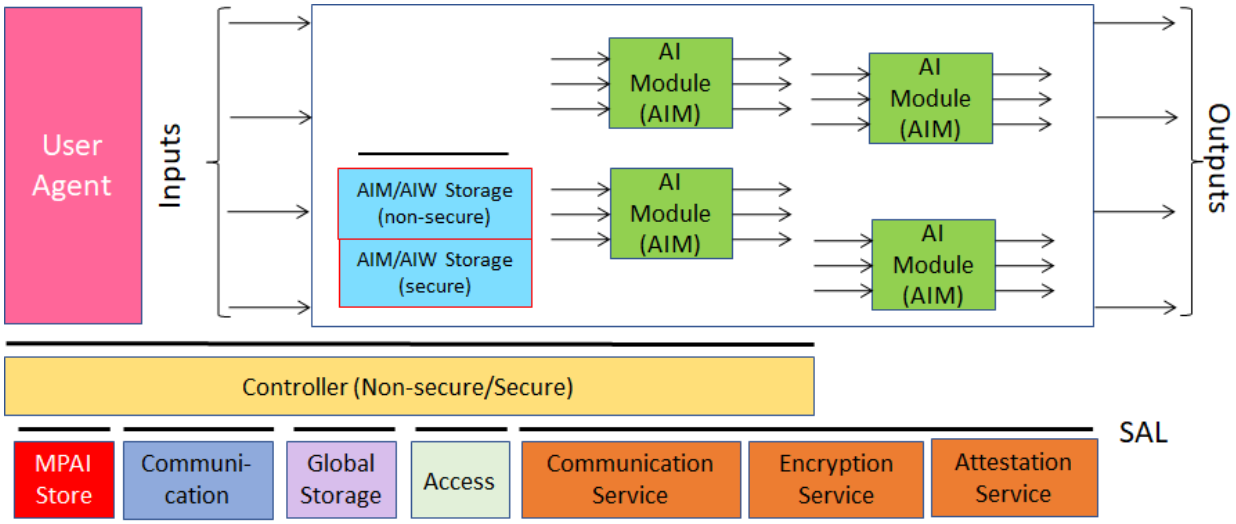

Figure 1 – The AI Framework (MPAI-AIF) V2 Reference Model

3. ***Technical Specification: Data Types, Formats, and Attributes (MPAI-TFA) V1.0*** specifies Qualifiers, a type of metadata supporting the operation of AIMs receiving data from other AIMs. Qualifiers convey information on Sub-Types (e.g., the type of colour), Formats (e.g., the type of compression and transport), and Attributes (e.g.,

semantic information in the Content). Although Qualifiers are human-readable, they are only intended to be used by AIMs. Therefore, Text, Speech, Audio, and Visual Data exchanged by AIWs and AIMs should be interpreted as being composed of Content (Text, Speech, Audio, and Visual as appropriate) and associated Qualifiers. The specifications of most MPAI Data Types reflect this point.

4. ***Technical Specification: Governance of the MPAI Ecosystem (MPAI-GME) V1.1*** defines the following elements:
    1. Standards, i.e., the ensemble of Technical Specifications, Reference Software, Conformance Testing, and Performance Assessment.
    2. Developers of MPAI-specified AIMs and Integrators of MPAI-specified AIWS (Implementers).
    3. MPAI Store in charge of making AIMs and AIWs submitted by Implementers available to Integrators and End Users.
    4. Performance Assessors, independent entities assessing the performance of implementations in terms of Reliability, Replicability, Robustness, and Fairness.
    5. End Users.

The interaction between and among actors of the MPAI Ecosystem are depicted in Figure 2.

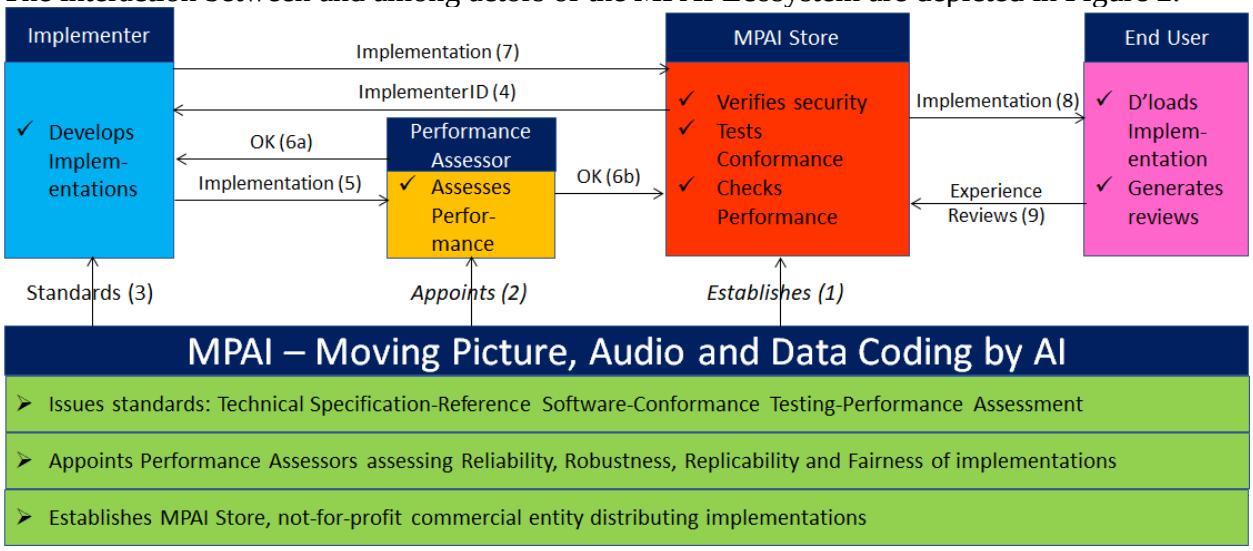

Figure 2 – The MPAI Ecosystem

2 Introduction (Informative)

Technical Specification: Context-based Audio Enhancement (MPAI-CAE) – Use Cases (CAE-USC) Version 2.2 (V2.2) collects use cases that improve the user experience of audio application using technologies that act on the input audio content using context information. The coverage of MPAI-CAE use cases includes entertainment, communication, teleconferencing, gaming, post-production, restoration etc. in a variety of contexts such as in the home, in the car, on-the-go, in the studio etc.

The technologies are specified as AI Workflows, AI Modules, and Data Types.

The currently specified use cases are *Emotion Enhanced Speech (EES)*, *Audio Recording Preservation (ARP)*, *Speech Restoration System (SSR)*, and *Enhanced Audioconference Experience (EAE)*.

3 Scope

Technical Specification: Context-based Audio Enhancement (MPAI-CAE) – Use Cases (CAE-USC) Version 2.2 (V2.2) specifies AI Workflows, AI Modules, and Data Types that support four use cases: *Emotion Enhanced Speech (EES), Audio Recording Preservation (ARP), Speech Restoration System (SSR),* and *Enhanced Audioconference Experience (EAE) and* the *Audio Scene Description (ASD)* Composite AIM.

Each Use Case normatively defines:

1. The Functions of the AIW and of the AIMs.
2. The Connections between and among the AIMs.
3. The Semantics and the Formats of the input and output data of the AIWs and its AIMs.

The word *normatively* implies an Implementation claiming Conformance, as follows:

1. An *AIW*, shall:
    1. Have the AIW Function specified in the relevant Use Case.
    2. Have all its AIMs and Connections conforming with the Reference Model of the AIW implementing the Use Case.
    3. Use Data Types specified at the relevant web page.
2. An *AIM*, shall:
    1. Have the AIM Function specified at the relevant web page.
    2. Use Data Types specified at the relevant web page.

Users of this Technical Specification should note that:

1. Implementers may use the Reference Software Implementations at the specified conditions.
2. The Conformance Testing specification can be used to test the Conformity of an Implementation to this Technical Specification.
3. Performance Assessors can assess the level of Performance of an Implementation based on the Performance Assessment specification of this Technical Specification.
4. Technical Specification: Governance of the MPAI Ecosystem (MPAI-GME) V1.1 specifies the operation of the MPAI Ecosystem.
5. Implementers and Users should consider the Notices and Disclaimers.

This version of the MPAI-CAE Technical Specification has been developed by the CAE-DC Development Committee. Future Versions may revise and/or extend the Scope of this Technical Specification.

4 Definitions

The Terms used in this standard whose first letter is capital have the meaning defined in *Table 1*. The general MPAI Terms are defined in *Table 2*.

Table 1 – Table of terms and definitions

Term	**Definition**
Access Copy Files	Set of files providing the information stored in an audio tape recording, including Restored Audio Files, suitable for audio information access, but not for long-term preservation.
Audio	Digital representation of an analogue audio signal sampled at a frequency between 8-192 kHz with a number of bits/sample between 8 and 64.
Audio Block	A set of consecutive Audio samples.
Audio Channel	A sequence of Audio Blocks.
Audio File	A .wav file [10].
Audio Object	Audio source which is in the audible frequency band.

Audio Scene Geometry	Spatial information for the Audio Objects which are included in an audio scene.
Audio Segment	An Audio Block with Start Time and an End Time Labels corresponding to the time of the first and last sample of the Audio Segment, respectively.
Audio-Visual File	A file containing audio and video according to the MP4 File Format [14].
Capstan	The capstan is a rotating spindle used to move recording tape through the mechanism of a tape recorder.
Damaged List	A list of strings of Texts corresponding to the Damaged Segments (if any) requiring replacement with synthetic segments.
Damaged Section	An Audio Segment which is damaged in its entirety and is contained in a Damaged Segment.
Damaged Segment	An Audio Segment containing only speech (and not containing music or other sounds) which is either damaged in its entirety or contains one or more Damaged Sections specified in the Damaged List.
Degree	Strength of a feature, specifically, with respect to Emotion, “High,” “Medium,” or “Low.”
Editing List	The description of the speed, equalisation and reading backwards corrections occurred during the restoration process.
Emotion	A Data Type representing the internal status of a human or avatar resulting from their interaction with the context or subsets of it, such as “Angry”, and “Sad”.
Emotionless Speech	An Audio File containing speech without music and other sounds, and in which little or no identifiable emotion is perceptible by native listeners.
Irregularity	An event of interest to preservation in *Table 26* and *Table 27*
Irregularity File	A JSON file containing information about Irregularities of the ARP inputs.
Irregularity Image	An image corresponding to an Irregularity.
JSON	JavaScript object notation [18].
Microphone Array Geometry	Description of the position of each microphone comprising the microphone array and specific characteristics such as microphone type, look directions, and the array type.
Model Utterance	An Audio Segment used as a model or demonstration of the Emotion to be added to Emotionless Speech in order to produce Speech with Emotion.
Multichannel Audio	A data structure containing at least 2 time-aligned interleaved Audio Channels.
Multichannel Audio Stream	A data structure containing Audio Objects packaged with Audio Scene Geometry.
Neural Network Speech Model	A Neural Network Model trained on Speech Segments for Modelling and used to synthesize replacements for the entire Damaged Segment or Damaged Sections within it.
Passthrough AIM	An AIM with the same input and output data of an AIM without executing the Function of that AIM. E.g., a Noise Cancellation AIM that does not cancel the noise.
Preservation Audio File	The input Audio File resulting from the digitisation of an audio open-reel tape to be preserved and, in case, restored.

Preservation Audio-Visual File	The input Audio-Visual File produced by a camera pointed to the playback head of the magnetic tape recorder and the synchronised Audio resulting from the tape digitisation process.
Preservation Image	A Video frame extracted from Preservation Audio-Visual File.
Preservation Master Files	Set of files providing the information stored in an audio tape recording without any restoration. As soon as the original analogue recordings is no more accessible, it becomes the new item for long-term preservation.
Restored Audio Files	Set of Audio Files derived from the Preservation Audio File, where potential speed, equalisation or reading backwards errors that occurred in the digitisation process have been corrected.
Restored Speech Segment	An Audio Segment in which the entire segment has been replaced by a synthetic speech segment, or in which each Damaged Segment has been replaced by a synthetic speech segment.
Speech Features	Descriptor representing a variety of information elements incorporated in a Speech Segment, e.g., personal identity, Personal Status, additional factors such as vocal tension, creakiness, whispery quality, etc.
Speech Segments for Modelling	A set of Audio Files containing speech segments used to train the Neural Network Speech Model.
Speech With Emotion File	An Audio File containing speech with emotional features.
Spherical Coordinate System	A coordinate system where the position of a point is specified by three numbers: the radial distance of that point from a fixed origin, its polar angle measured from a fixed zenith direction, and the azimuthal angle of its orthogonal projection on a reference plane.
Spherical Grid Resolution	The maximum spherical angle between any two neighbouring sampled points on a sphere.
Text List	List of texts to be converted into speech by the Speech Synthesis for Restoration AIM.
Time Code	Number of ms from 1970-01-01T00:00:00.000 according to [8].
Time Label	A measure of time from a context-dependent zero time expressed as HH:mm:ss.SSS.
Transform Audio	A frequency representation of Audio.
Enhanced Transform Audio	Transform Audio whose samples are Enhanced Transform Audio samples.
Useful Signal	Digital signal resulting from the A/D conversion of the analogue signal recorded in an audio tape.

Table 2 – MPAI-wide Terms

Term	**Definition**
Access	Static or slowly changing data that are required by an application such as domain knowledge data, data models, etc.
AI Framework (AIF)	The environment where AIWs are executed.
AI Workflow (AIW)	An organised aggregation of AIMs implementing a Use Case receiving AIM-specific Inputs and producing AIM-specific Outputs according to its Function.
AI Module (AIM)	A processing element receiving AIM-specific Inputs and producing AIM-specific Outputs according to according to its Function.
Application Standard	An MPAI Standard designed to enable a particular application domain.

Channel	A connection between an output port of an AIM and an input port of an AIM. The term “connection” is also used as synonymous.
Communication	The infrastructure that implements message passing between AIMs.
Composite AIM	An AIM aggregating more than one AIM.
Component	One of the 7 AIF elements: Access, Communication, Controller, Internal Storage, Global Storage, MPAI Store, and User Agent.
Composite AIM	
Conformance	The attribute of an Implementation of being a correct technical Implementation of a Technical Specification.
Conformance Tester	An entity authorised by MPAI to Test the Conformance of an Implementation.
Conformance Testing	The normative document specifying the Means to Test the Conformance of an Implementation.
Conformance Testing Means	Procedures, tools, data sets and/or data set characteristics to Test the Conformance of an Implementation.
Connection	A channel connecting an output port of an AIM and an input port of an AIM.
Controller	A Component that manages and controls the AIMs in the AIF, so that they execute in the correct order and at the time when they are needed.
Data Format	The standard digital representation of data.
Data Semantics	The meaning of data.
Ecosystem	The ensemble of the following actors: MPAI, MPAI Store, Implementers, Conformance Testers, Performance Testers and Users of MPAI-AIF Implementations as needed to enable an Interoperability Level.
Explainability	The ability to trace the output of an Implementation back to the inputs that have produced it.
Fairness	The attribute of an Implementation whose extent of applicability can be assessed by making the training set and/or network open to testing for bias and unanticipated results.
Function	The operations effected by an AIW or an AIM on input data.
Global Storage	A Component to store data shared by AIMs.
Internal Storage	A Component to store data of the individual AIMs.
Identifier	A name that uniquely identifies an Implementation.
Implementation	
Implementer	A legal entity implementing MPAI Technical Specifications.
ImplementerID (IID)	A unique name assigned by the ImplementerID Registration Authority to an Implementer.
ImplementerID Registration Authority (IIDRA)	The function within the MPAI Store to assign ImplementerID’s to Implementers.
Interoperability	The ability to functionally replace an AIM with another AIM having the same Interoperability Level.
Interoperability Level	The attribute of an AIW and its AIMs to be executable in an AIF Implementation and to:
Knowledge Base	Structured and/or unstructured information made accessible to AIMs via MPAI-specified interfaces.
Message	A sequence of Records transported by Communication through Channels.

Normativity	The set of attributes of a technology or a set of technologies specified by the applicable parts of an MPAI standard.
Performance	The attribute of an Implementation of being Reliable, Robust, Fair and Replicable.
Performance Assessment	The normative document specifying the procedures, the tools, the data sets and/or the data set characteristics to Assess the Grade of Performance of an Implementation.
Performance Assessment Means	Procedures, tools, data sets and/or data set characteristics to Assess the Performance of an Implementation.
Performance Assessor	An entity authorised by MPAI to Assess the Performance of an Implementation in a given Application domain.
Profile	A particular subset of the technologies used in MPAI-AIF or an AIW of an Application Standard and, where applicable, the classes, other subsets, options and parameters relevant to that subset.
Record	A data structure with a specified structure.
Reference Model	The AIMs and theirs Connections in an AIW.
Reference Software	A technically correct software implementation of a Technical Specification containing source code, or source and compiled code.
Reliability	The attribute of an Implementation that performs as specified by the Application Standard, profile and version the Implementation refers to, e.g., within the application scope, stated limitations, and for the period of time specified by the Implementer.
Replicability	The attribute of an Implementation whose Performance, as Assessed by a Performance Assessor, can be replicated, within an agreed level, by another Performance Assessor.
Robustness	The attribute of an Implementation that copes with data outside of the stated application scope with an estimated degree of confidence.
Service Provider	An entrepreneur who offers an Implementation as a service (e.g., a recommendation service) to Users.
Standard	The ensemble of Technical Specification, Reference Software, Conformance Testing and Performance Assessment of an MPAI application Standard.
Technical Specification	(Framework) the normative specification of the AIF. (Application) the normative specification of the set of AIWs belonging to an application domain along with the AIMs required to Implement the AIWs that includes:
Testing Laboratory	A laboratory accredited by MPAI to Assess the Grade of Performance of Implementations.
Time Base	The protocol specifying how Components can access timing information.
Topology	The set of AIM Connections of an AIW.
Use Case	A particular instance of the Application domain target of an Application Standard.
User	A user of an Implementation.
User Agent	The Component interfacing the user with an AIF through the Controller.
Version	A revision or extension of a Standard or of one of its elements.
Zero Trust	A model of cybersecurity primarily focused on data and service protection that assumes no implicit trust.

5 References

5.1 Normative References

This standard normatively references the following technical specifications, both from MPAI and other standard organisations:

1. 1. MPAI; Technical Specification: The governance of the MPAI Ecosystem (MPAI-GME) V1.1; https://mpai.community/standards/mpai-gme/
2. 2. MPAI; Technical Specification: Artificial Intelligence Framework (MPAI-AIF) V2.0; https://mpai.community/standards/mpai-aif/
3. 3. MPAI; Technical Specification: Connected Autonomous Vehicles (MPAI-CAV) – Architecture V1.0; https://mpai.community/standards/mpai-cav/
4. 4. MPAI; Technical Specification: Multimodal Conversation (MPAI-MMC) V2.1; https://mpai.community/standards/mpai-mmc/
5. 5. MPAI; Technical Specification: Object and Scene Description (MPAI-OSD); https://mpai.community/standards/mpai-osd/
6. 6. MPAI; Technical Specification: Portable Avatar Format (MPAI-PAF); https://mpai.community/standards/mpai-paf/
7. 7. A Universally Unique IDentifier (UUID) URN Namespace; IETF RFC 4122; July 2005.
8. 8. Date and Time on the Internet: Time Stamps; IETF RFC 3339; July 2002.
9. 9. Universal Coded Character Set (UCS): ISO/IEC 10646; December 2020.
10. ITU-R BS.2088-1 (10/2019) – Long-form file format for the international exchange of audio programme materials with metadata.
11. ISO/IEC 14496-10; Information technology – Coding of audio-visual objects – Part 10: Advanced Video Coding.
12. ISO/IEC 23008-2; Information technology – High efficiency coding and media delivery in heterogeneous environments – Part 2: High Efficiency Video Coding.
13. ISO/IEC 23094-1; Information technology – General video coding – Part 1: Essential Video Coding.
14. ISO/IEC 14496-12; Information technology – Coding of audio-visual objects – Part 12: ISO base media file format.
15. ZIP format, https://pkware.cachefly.net/webdocs/casestudies/APPNOTE.TXT.
16. Neural Network Exchange Format; https://www.khronos.org/registry/NNEF/specs/1.0/nnef-1.0.4.pdf; Khronos.
17. Open Neural Network Exchange (ONNX) format; https://www.ONNX.ai.
18. The JavaScript Object Notation (JSON) Data Interchange Format; https://datatracker.ietf.org/doc/html/rfc8259; IETF rfc8259; December 2017.
19. BS EN 60094-1:1994, BS 6288-1: 1994, IEC 94-1:1981 – Magnetic tape sound recording and reproducing systems – Part 1: Specification for general conditions and requirements.
20. K. Bradley, IASA TC-04 Guidelines in the Production and Preservation of Digital Audio Objects: standards, recommended practices, and strategies., 2nd ed. International Association of Sound and Audiovisual Archives, (2009): 2014.
21. MPAI; The MPAI Statutes; https://mpai.community/statutes/
22. MPAI; The MPAI Patent Policy; https://mpai.community/about/the-mpai-patent-policy/.
23. Framework Licence of the Context-based Audio Enhancement Technical Specification (MPAI-CAE); https://mpai.community/standards/mpai-cae/framework-licence/
24. ITU-R BS.2088-1: Long-form file format for the international exchange of audio programme materials with metadata.

25. ITU-T T-81: Information technology — Digital compression and coding of continuous-tone still images: Requirements and guidelines.

5.2 Informative References

The references provided here are for information purpose.

26. Ekman, Paul (1999), "Basic Emotions", in Dalgleish, T; Power, M (eds.), Handbook of Cognition and Emotion (PDF), Sussex, UK: John Wiley & Sons.
27. B. Rafaely, Fundamentals of spherical array processing, Springer, 2018.

6 AI Workflows

Technical Specification: Context-based Audio Enhancement (CAE-USC) V2.2 specifies the following AI Workflows:

Emotion-Enhanced Speech	Audio Recording Preservation
Speech Restoration System	Enhanced Audioconference Experience

Note: The specification of Audio Recording Preservation and Enhanced Videoconference Experience is unchanged from V2.1.

Each AI Workflow includes the specification of:

1. Functions of the AI Workflow
2. Reference Model of the AI Workflow
3. I/O data of the AI Workflow
4. Functions of AI Modules
5. I/O data of the AI Modules
6. AIW, AIMs, and JSON Metadata

Each AIW, AIM, and Data Type is specified by individual web pages that can be referenced by other Technical Specifications. The list of links to all AIWs, AIMs, and Data Types is available.

7 AI Modules

Table 1 provides web links to the specification of AI Modules and JSON Metadata used and specified by CAE-USC V2.2. The first columns indicates which CAE-USC Version has specified the AIM.

Table 1 – AI Modules used and specified by CAE-USC V2.2

V	Acronym	AIM Name	JSON
2.1	MMC-AAP	**Audio Analysis for Preservation**	**File**
2.2	MMC-AAT	**Audio Analysis Transform**	**File**
2.2	MMC-ABS	**Audio Basic Scene Description**	**File**
2.1	MMC-ADP	**Audio Description Packaging**	**File**
2.2	MMC-AMX	**Audio Descriptors Multiplexing**	**File**
2.2	MMC-AOI	**Audio Object Identification**	**File**
2.2	MMC-ASD	**Audio Scene Description**	**File**
2.2	MMC-ASE	**Audio Separation and Enhancement**	**File**
2.2	MMC-ASL	**Audio Source Localisation**	**File**
2.2	MMC-AST	**Audio Synthesis Transform**	**File**
2.2	MMC-EFP	**Emotion Feature Production**	**File**
2.1	MMC-NCM	**Noise Cancellation Module**	**File**
2.2	MMC-NEI	**Neural Emotion Insertion**	**File**

2.1	MMC-PAP	**Packaging for Audio Preservation**	**File**
2.2	MMC-PEI	**Prosodic Emotion Insertion**	**File**
2.1	MMC-SDS	**Speech Detection and Separation**	**File**
2.2	MMC-SF1	**Speech Feature Analysis 1**	**File**
2.2	MMC-SF2	**Speech Feature Analysis 2**	**File**
2.1	MMC-SFD	**Sound Field Description**	**File**
2.2	MMC-SMC	**Speech Model Creation**	**File**
2.2	MMC-SRA	**Speech Restoration Assembly**	**File**
2.2	MMC-SSR	**Speech Synthesis for Restoration**	**File**
2.1	MMC-TAR	**Tape Audio Restoration**	**File**
2.1	MMC-TIC	**Tape Irregularity Classification**	**File**
2.1	MMC-VAP	**Video Analysis for Preservation**	**File**

8 Data Types

The Table provides web links to the specification of Data Types used by CAE-TEC V2.2 and not specified by MPAI-CAE V2.1. The complete list of Data Types specified by V2.1 and V2.2 is available.

Audio Basic Scene Descriptors	**Audio Basic Scene Geometry**	**Audio Object**	**Audio Scene Descriptors**
Audio Scene Geometry	**Damaged List**	**Microphone Array Geometry**	

9 Informative Examples

9.1 Audio Scene Geometry

An example of Audio Scene Geometry.


```
{
 “BlockIndex”: 1,
 “BlockStart”: 1631536788000,
 “BlockEnd”: 1631536788063,
 “SpeechCount”: 2,
 “SpeechList”: [
        {
            “SpeechID”: “09859d16-3c73-4bb0-9c74-91b451e34925”,
            “ChannelID”: 1,
            “AzimuthDirection”: 90.0,
            “ElevationDirection”: 30.0,
            “Distance”: 2.0,
            “DistanceFlag”: false
        },
        {
            “SpeechID”: “3cdc2973-e95e-4125-acb7-121ad89067ef”,
            “ChannelID”: 2,
            “AzimuthDirection”: 180.0,
            “ElevationDirection”: 30.0,
            “Distance”: 1.27,
            “DistanceFlag”: false
```

```
        }
 ],
 “SourceDetectionMask”: [0,1]
}
```

9.2 Damaged List

An example of a damaged list JSON file:

```
{
     “DamagedSections”: [
   {
     “SegmentStart”: “00:00:01.351”,
      “SegmentEnd”: “00:01:55.654”,
    },
    {
     “SegmentStart”: “00:01:55.654”,
     “SegmentEnd”: “00:02:35.168”,
    }
     ]
}
```

9.3 . Editing List

Example of a complete Editing List with two elements: the first related to reading backwards error, whereas the second to speed and equalisation errors.

```
{
   “OriginalSpeedStandard”: 15,
   “OriginalEqualisationStandard”: “IEC1”,
   “OriginalSampleFrequency”: 96000,
   “Restorations”: [{
     “RestorationID”: “09859d16-3c73-4bb0-9c74-91b451e34925”,
     “PreservationAudioFileStart”: “00:00:00.000”,
     “PreservationAudioFileEnd”: “00:00:05.125”,
     “RestoredAudioFileURI”: “http://www.place_to_be_defined.com/restored_1”,
     “ReadingBackwords”: true,
     “AppliedSpeedStandard”: 15,
     “AppliedSampleFrequency”: 96000,
     “OriginalEqualisationStandard”: “IEC1”
   },
   {
     “RestorationID”: “3cdc2973-e95e-4125-acb7-121ad89067ef “,
     “PreservationAudioFileStart”: “00:00:05.125”,
     “PreservationAudioFileEnd”: “00:00:15.230”,
     “RestoredAudioFileURI”: “http://www.place_to_be_defined.com/restored_2”,
     “ReadingBackwords”: false,
     “AppliedSpeedStandard”: 7.5,
     “AppliedSampleFrequency”: 48000,
     “OriginalEqualisationStandard”: “IEC2”
   }]
}
```

9.4 Irregularity File

An example of Irregularity File from Audio Analyser to Video Analyser is:


```
{
   “Offset”: 150,
   “Irregularities”: [{
      “IrregularityID”: “09859d16-3c73-4bb0-9c74-91b451e34925”,
      “Source”: “a”,
      “TimeLabel”: “00:02:45.040”
   },{
      “IrregularityID”: “3cdc2973-e95e-4125-acb7-121ad89067ef”,
      “Source”: “a”,
      “TimeLabel”: “00:04:89.020”
   }]
}
```


An example of Irregularity File from Video Analyser to Audio Analyser is:


```
{
   “Irregularities”: [{
      “IrregularityID”: “09859d16-3c73-4bb0-9c74-91b451e34925”,
      “Source”: “v”,
      “TimeLabel”: “00:02:45.040”
   },{
      “IrregularityID”: “3cdc2973-e95e-4125-acb7-121ad89067ef”,
      “Source”: “v”,
      “TimeLabel”: “00:04:89.020”
   }]
}
```


An example of Irregularity File from Audio Analyser to Tape Irregularity Classifier is:


```
{
   “Offset”: 150,
   “Irregularities”: [{
      “IrregularityID”: “09859d16-3c73-4bb0-9c74-91b451e34925”,
      “Source”: “a”,
      “TimeLabel”: “00:02:45.040”,
      “AudioSegmentURI”: “http://www.place_to_be_defined.com/audio_segment_1”,
      “IrregularityType”: “ssv”,
      “IrregularityProperties: {
         “ReadingSpeedStandard”: 15,
         “ReadingEqualisationStandard”: “IEC1”,
         “WritingSpeedStandard”: 7.5,
         “WritingEqualisationStandard”: “IEC2”
      }
   },{
      “IrregularityID”: “3cdc2973-e95e-4125-acb7-121ad89067ef”,
      “Source”: “v”,
      “TimeLabel”: “00:04:89.020”,
      “AudioSegmentURI”: “http://www.place_to_be_defined.com/audio_segment_2”
   }]
}
```


An example of Irregularity File from Video Analyser to Tape Irregularity Classifier is:

```
{
  “Offset”: 150,
  “Irregularities”: [{
    “IrregularityID”: “09859d16-3c73-4bb0-9c74-91b451e34925”,
    “Source”: “a”,
    “TimeLabel”: “00:02:45.040”,
    “ImageURI”: “http://www.place_to_be_defined.com/image_1”
  },{
    “IrregularityID”: “3cdc2973-e95e-4125-acb7-121ad89067ef”,
    “Source”: “v”,
    “TimeLabel”: “00:04:89.020”,
    “ImageURI”: “http://www.place_to_be_defined.com/image_2”
  }]
}
```


An example of Irregularity File from Tape Irregularity Classifier to Tape Audio Restoration is:


```
{
  “Irregularities”: [{
    “IrregulatityID”: “09859d16-3c73-4bb0-9c74-91b451e34925”,
    “Source”: “a”,
    “TimeLabel”: “00:02:45.040”,
    “IrregularityType”: “ssv”,
    “IrregularityProperties: {
      “ReadingSpeedStandard”: 15,
      “ReadingEqualisationStandard”: “IEC1”,
      “WritingSpeedStandard”: 7.5,
      “WritingEqualisationStandard”: “IEC2”
    }
  },{
    “IrregulatityID”: “3cdc2973-e95e-4125-acb7-121ad89067ef”,
    “Source”: “a”,
    “TimeLabel”: “00:04:89.020”,
    “IrregularityType”: “esv”,
    “IrregularityProperties: {
      “ReadingSpeedStandard”: 7.5,
      “ReadingEqualisationStandard”: “IEC2”,
      “WritingSpeedStandard”: 7.5,
      “WritingEqualisationStandard”: “IEC1”
    }
  }]
}
```


An example of Irregularity File from Tape Irregularity Classifier to Packager is:


```
{
  “Offset”: 150,
  “Irregularities”: [{
    “IrregulatityID”: “09859d16-3c73-4bb0-9c74-91b451e34925”,
    “Source”: “v”,
    “TimeLabel”: “00:02:45.040”,
    “IrregularityType”: “sot”,
    “ImageURI”: “http://www.place_to_be_defined.com/image_1”
  },{
```

```
      “IrregulatityID”: “3cdc2973-e95e-4125-acb7-121ad89067ef”,
      “Source”: “b”,
      “TimeLabel”: “00:04:89.020”,
      “IrregularityType”: “sp”,
      “ImageURI”: “http://www.place_to_be_defined.com/image_2”
   }]
}
```


9.5 Microphone Array Geometry


```
{
 “MicrophoneArrayType”: 0,
 “MicrophoneArrayScat”: 0,
 “MicrophoneArrayFilterURI”: “https://mpai.community/standards/mpai-cae/”,
 “SamplingRate”: 4,
 “SampleType”: 0,
 “BlockSize”: 3,
 “NumberofMicrophones”: 4,
 “MicrophoneList”: [
        {
            “xCoord”: 1.0,
            “yCoord”: 2.0,
            “zCoord”: 3.0,
            “directivity”: 0,
            “micxLookCoord”: 70.2,
            “micyLookCoord”: 75.5,
            “miczLookCoord”: 87.3
        },
        {
            “xCoord”: 5.3,
            “yCoord”: 5.6,
            “zCoord”: 74.3,
            “directivity”: 1,
            “micxLookCoord”: 67.9,
            “micyLookCoord”: 75.2,
            “miczLookCoord”: 90.0
        },
        {
            “xCoord”: 34.2,
            “yCoord”: 65.2,
            “zCoord”: 56.9,
            “directivity”: 2,
            “micxLookCoord”: 56.8,
            “micyLookCoord”: 87.9,
            “miczLookCoord”: 78.3
        },
        {
            “xCoord”: 34.9,
            “yCoord”: 29.7,
            “zCoord”: 89.8,
            “directivity”: 3,
```

```
                "micxLookCoord": 56.9,
                "micyLookCoord": 65.4,
                "miczLookCoord": 72.9
            }
  ],
  "MicrophoneArrayLookCoord": [{
    "xLookCoord": 56.0,
    "yLookCoord": 90.0,
    "zLookCoord": 86.3
  }]
}
```

9.6 Prosodic Speech Features

```
{
   "intonations": [{
      "pitch": 300,
      "intensity": 88.7,
      "duration":100.0
   },{
      "pitch": 180,
      "intensity": 85.2,
      "duration":98.0
   },{
      "pitch": 280,
      "intensity": 92.5,
      "duration":92.0
   },{
      "pitch": 230,
      "intensity": 81.9,
      "duration":98.0
   },{
      "pitch": 150,
      "intensity": 78.3,
      "duration":98.0
   }],
   "unit": "phoneme"
}
```

9.7 Neural Speech Features

```
[
   1.456,
   5.1289,
   0.12,
   12345.54378,
   12389943.2837,
   58.29
]
```

10 Notices and Disclaimers

The notices and legal disclaimers given below shall be borne in mind when downloading and using approved MPAI Standards.

In the following, "Standard" means the collection of four MPAI-approved and published documents: "Technical Specification", "Reference Software" and "Conformance Testing" and, where applicable, "Performance Testing".

Life cycle of MPAI Standards

MPAI Standards are developed in accordance with the MPAI Statutes. An MPAI Standard may only be developed when a Framework Licence has been adopted. MPAI Standards are developed by especially established MPAI Development Committees who operate on the basis of consensus, as specified in Annex 1 of the MPAI Statutes. While the MPAI General Assembly and the Board of Directors administer the process of the said Annex 1, MPAI does not independently evaluate, test, or verify the accuracy of any of the information or the suitability of any of the technology choices made in its Standards.

MPAI Standards may be modified at any time by corrigenda or new editions. A new edition, however, may not necessarily replace an existing MPAI standard. Visit the web page to determine the status of any given published MPAI Standard.

Description on MPAI Standards are welcome from any interested parties, whether MPAI members or not. Comments shall mandatorily include the name and the version of the MPAI Standard and, if applicable, the specific page or line the comment applies to. Comments should be sent to the MPAI Secretariat. Comments will be reviewed by the appropriate committee for their technical relevance. However, MPAI does not provide interpretation, consulting information, or advice on MPAI Standards. Interested parties are invited to join MPAI so that they can attend the relevant Development Committees.

Coverage and Applicability of MPAI Standards

MPAI makes no warranties or representations concerning its Standards, and expressly disclaims all warranties, expressed or implied, concerning any of its Standards, including but not limited to the warranties of merchantability, fitness for a particular purpose, non-infringement etc. MPAI Standards are supplied "AS IS".

The existence of an MPAI Standard does not imply that there are no other ways to produce and distribute products and services in the scope of the Standard. Technical progress may render the technologies included in the MPAI Standard obsolete by the time the Standard is used, especially in a field as dynamic as AI. Therefore, those looking for standards in the Data Compression by Artificial Intelligence area should carefully assess the suitability of MPAI Standards for their needs.

IN NO EVENT SHALL MPAI BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO: THE NEED TO PROCURE 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 PUBLICATION, USE OF, OR RELIANCE UPON ANY STANDARD, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE AND REGARDLESS OF WHETHER SUCH DAMAGE WAS FORESEEABLE.

MPAI alerts users that practicing its Standards may infringe patents and other rights of third parties. Submitters of technologies to this standard have agreed to licence their Intellectual Property according to their respective Framework Licences.

Users of MPAI Standards should consider all applicable laws and regulations when using an MPAI Standard. The validity of Conformance Testing is strictly technical and refers to the correct implementation of the MPAI Standard. Moreover, positive Performance Assessment of an implementation applies exclusively in the context of the MPAI Governance and does not imply

compliance with any regulatory requirements in the context of any jurisdiction. Therefore, it is the responsibility of the MPAI Standard implementer to observe or refer to the applicable regulatory requirements. By publishing an MPAI Standard, MPAI does not intend to promote actions that are not in compliance with applicable laws, and the Standard shall not be construed as doing so. In particular, users should evaluate MPAI Standards from the viewpoint of data privacy and data ownership in the context of their jurisdictions.

Implementers and users of MPAI Standards documents are responsible for determining and complying with all appropriate safety, security, environmental and health and all applicable laws and regulations.