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

|  |  |
| --- | --- |
| **N549** | 2022/02/23 |
| **Source** | CAE-DC |
| **Title** | Introduction to MPAI-CAE |
| **Target** | MPAI Members |

The Moving Picture, Audio and Data Coding by Artificial Intelligence ([MPAI](https://mpai.community/)) community is an international, unaffiliated, non-profit organisation having the mission to 1) develop standards for a) Artificial Intelligence (AI)-based data coding and b) integration of data coding components into systems, and 2) to attach clear Intellectual Property Rights licensing frameworks to its standards.

AI promises to offer superior efficiency in data coding – for example, for data compression or feature-based description – as compared with other current coding technologies.

The scope of the MPAI [*Context-based Audio Enhancement*](https://mpai.community/standards/mpai-cae/) *(MPAI-CAE)* application standard is to improve the user experience for audio-related applications including 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. using context information to act on the input audio content, and potentially deliver the processed output via an appropriate protocol.

Four use cases have been identified and standard technologies provided in Version 1 of the standard:

1. *Emotion Enhanced Speech (EES)* enables a user to convert an individual emotionless speech segment to a segment that has a specified emotion expressed either as a tag belonging to a standard list of emotions or derived by extracting features from a model utterance.
2. *Audio Recording Preservation (ARP),* enables a user to create of a copy of the digitised audio of an open reel magnetic tape for long-term preservation and of a restored access copy for correct play back.
3. *Speech Restoration System (SSR)* enables a user to restore a damaged segment of an audio segment containing speech from a single speaker by providing the text of damaged segment and neural network model obtained from undamaged parts of the input speech segment.
4. *Enhanced Audioconference Experience (EAE)*. enables a user to improve auditory experience in an audioconference by 1) extracting the speech signals from individual speakers, 2) reducing the background noise and the reverberation from a microphone array, and 3) extracting the spatial attributes of conference participants to enable spatial representation of speech signals at the receiver side.

The role of standards is to enable interoperability of implementations. By providing the means to verify interoperability, standards create ecosystems supporting the life cycle of specification, implementation, security verification, conformance testing, distribution and consumption.

Because of their particular nature, however, AI standard should also provide the means to assess the degree of reliability of an implementation.

The [*Governance of the MPAI Ecosystem*](https://mpai.community/standards/mpai-gwe/) *(MPAI-GME)* standard identifies 1) the MPAI Store as the actor in charge of security verification, conformance testing, and distribution of implementations and 2) the Performance Assessors as the actors in charge of assessing that implementation are reliable.

MPAI application standards are components-based in the sense that a full application, such as one of the MPAI-CAE use cases above, is typically implemented with a set of basic processing elements called *AI Modules* (AIMs) connected to form *AI Workflows* (AIWs) executed in an *AI Framework* (AIF)*.* The [*AI Framework (MPAI-AIF)*](https://mpai.community/standards/mpai-aif/) standard specifies architecture, interfaces, protocols and Application Programming Interfaces (API).

MPAI develops its standards through a rigorous process combining openness to all interested parties, including non-members (when requirements for a new standard are identified) and confidentiality (when technology employing a standard is integrated).

Diagram

Description automatically generated

*Figure 1 – The stages of the MPAI standards development process*