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

**Public Document**

|  |  |
| --- | --- |
| **N364** | 2021/09/30 |
| **Source** | Communication AC |
| **Title** | Introduction to MPAI-MMC |
| **Target** | MPAI Members |

The [Moving Picture, Audio and Data Coding by Artificial Intelligence (MPAI) community](https://mpai.community/) is an international, unaffiliated, non-profit organization. Its mission: to develop Artificial Intelligence (AI)-based data coding standards, with clear Intellectual Property Rights licensing frameworks. Research has established that AI-based technologies enable superior efficiency in data coding – for example, for data compression or feature-based description – as compared with other current coding technologies.

MPAI develops its standards through a rigorous process combining openness to all interested parties (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*

Like all MPAI application standards, MPAI Multimodal Communication (MPAI-MMC) is implemented within the *AI Framework* (AIF); and this framework, in turn, is specified by the *MPAI-AIF Standard*, whose *Reference Model* is depicted in *Figure 1*. AIFs exec­ute *AI Workflows* (AIWs), composed of basic processing elements called *AI Modules* (AIMs).

Diagram

Description automatically generated

*Figure 2 – The AI Framework (AIF) Reference Model and its Components*

The MPAI-MMC Application Standard presently specifies five use cases.

* In the Conversation with Emotion (CWE) use case, a human holds an audio-visual conver­sation with a computational system personified by a synthetic voice and an animated face.
* In the Multimodal Question Answering (MQA) use case, a human user requests and receives from a computational system spoken information concerning a displayed object.
* In three conversational translation use cases, a computational system translates from one spoken language to one or more other spoken languages. The translation path may be one-to-one from Language A to B only (in the Unidirectional Speech Translation (UST) use case); from Language A to B and vice versa (in the Bidirectional Speech Translation (BST) use case); or from Language A to B, C, … (in the One-to-Many Speech Translation (MST) use case). Synthetic spoken output can preserve specified features of the source language speech.

In all such use cases, MPAI-MMC specifies the implementation architecture; the AI modules comprising the topology; and the formats of the AIMs’ input and output data. And, based on the MPAI-AIF specification, it defines the workflow linking the AIMs, along with their metadata.

The AIMs specified by MPAI-MMC or comparable application standards – modules typically based on AI technologies but in some cases on data processing technologies – can be reused in more than one application scenario.

Multimodality is the essence of MPAI-MMC. Text, speech, and video are exploited, in both input and output data, to enhance user experience and improve human-machine interaction. This application standard will be valuable in AI industries aiming to enhance services based on human machine interaction of all sorts, particularly when the emphasis is on utilization of multimodal user interfaces enabling emotional expression through natural language and visual communication.

The MPAI-MMC application standard is *reliable* because it has been developed by industry experts who have assiduously contributed to the MPAI- MMC Technical Specification. The Reference Software, another specification of the MPAI- MMC Standards suite, is a conforming sample implemen­tation of the Standard, released with the MPAI [software licence](https://mpai.community/about/licence/) (a modified Berkeley Source Distribution or BSD-type licence).

Implementers conforming to a standard can upload their implementations to the MPAI Store. There the latter are tested for security and correct implementation via Conformance Testing, another specification of the MPAI- MMC standards suite.

The possibility of testing an implementation for conformance to a standard will provide unique benefits to the AI-based software industry. Testing performed by the MPAI Store ensures that implementations offered to consumers via the Store are interoperable with other implementations. Moreover, new and smarter AIMs with the same functionality can replace older and less efficient ones. Implementations conforming to MPAI standards are expected to exemplify the state of the AI art; however, given the intrinsic competition between AIM developers, the same standards also feed innovation. As new AIMs are added or substituted, the standard will gradually foster increasingly sophisticated and function-rich implementations.

MPAI-MMC will benefit society, as well. Consumers will receive guidance in selecting state-of- the-art and reliable applications, profiting from the judgments of Performance Assessors, who will rate AIMs from competing implementers. The Assessors, in turn, will be guided by Performance Assessment, the fourth specification of the MPAI- MMC suite.

[NOTE, however: only the MPAI-MMC Technical Specification is available at this time.]