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

MPAI Technical Specification

Data Types, Formats and Attributes (MPAI-TFA)

V1.2

Technical Specification
Data Types, Formats and Attributes (MPAI-TFA) – V1.2

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. *Obtain* conforming implementations possibly with a performance assessment report from a trusted source through the MPAI Store.

Today, the MPAI organisation operated 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) typically including interconnected AI Modules (AIM). MPAI-AIF supports small- and large-scale high-performance components and promotes solutions with improved explainability.
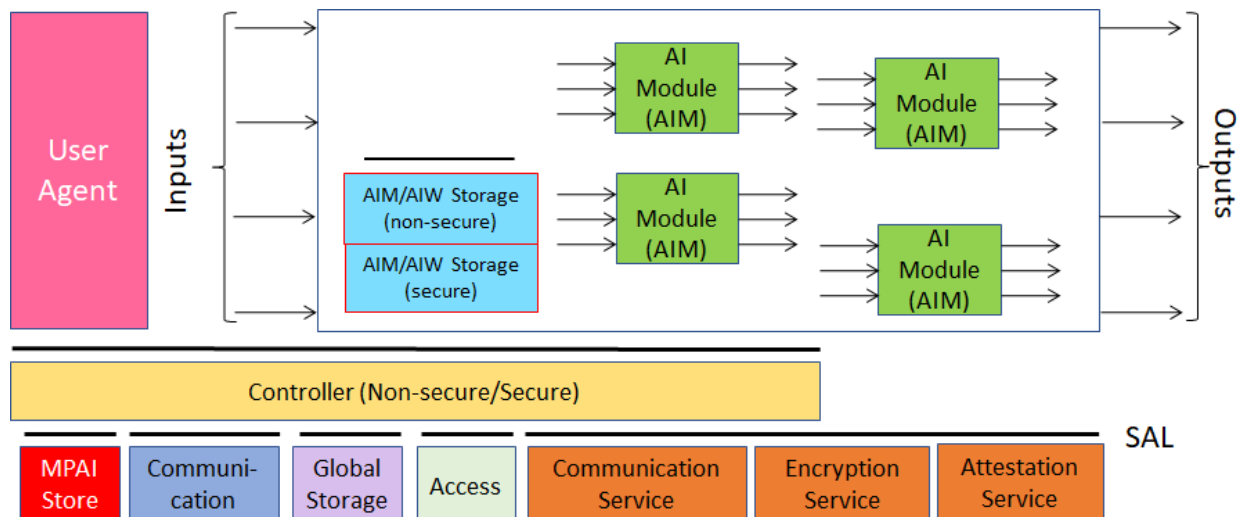

Figure 1 – The AI Framework (MPAI-AIF) V2 Reference Model

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, Visual, and other Data exchanged by AIWs and AIMs should be interpreted as being composed of Content (Text, Speech, Audio, and Visual as appropriate) and associated Qualifiers. Therefore a Text Object is composed of Text Data and Text Qualifier. The specification of most MPAI Data Types reflects 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. <u>MPAI Store</u> in charge of making AIMs and AIWs submitted by Implementers available to Integrators and End Users.
4. <u>Performance Assessors</u>, independent entities assessing the performance of implementations in terms of Reliability, Replicability, Robustness, and Fairness.
5. <u>End Users</u>.

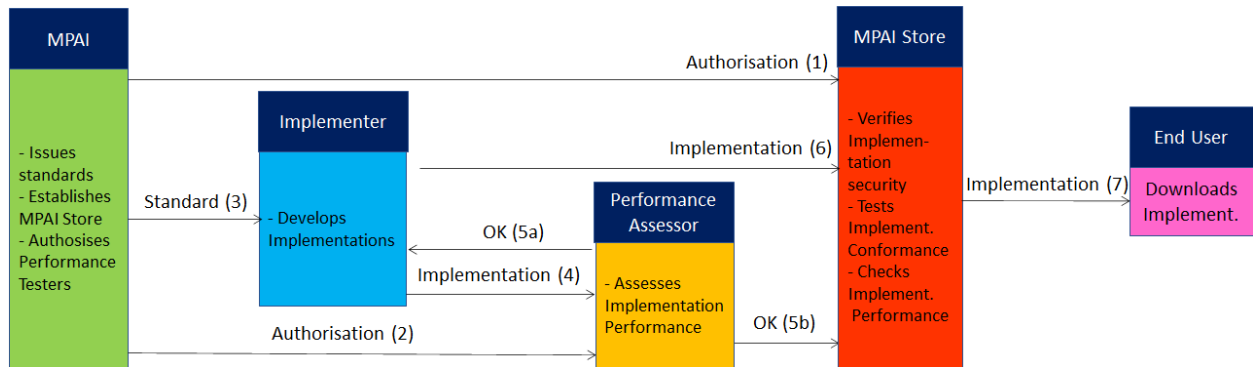The interaction between and among actors of the MPAI Ecosystem are depicted in Figure 2.

Figure 2 – The MPAI Ecosystem

2 Introduction (informative)

Technical Specification: <u>AI Framework</u> (**MPAI-AIF**) **V2.1** is a key element of the MPAI approach to AI-based Data Coding standards. It is based on a framework enabling initialisation, dynamic configuration, and control of AIWs in the standard AI Framework environment depicted in <u>Figure 1</u>. The Data Data produced by executing specific functions by AI Modules (AIM) are communicate to other AIMs in the AIW.

The effectiveness of the functions perform may the AIMs may be improved if they know more about the **capabilities of the AIMs** they are connected to and the **Data** they receive. For instance, an instance of the MPAI <u>Natural Language Processing</u> (MMC-NLU) AIM has the task to refine the text received and produce its <u>Meaning</u> using three levels of information:

1. Just the **input text**.
2. The **object identifiers** referenced in the text.
3. The **object context** in a relevant space.

The cases correspond with different levels of AIM capabilities that likely correspond to improved accuracy of the refined text and Meaning produced when moving from the first to the third case.

Technical Specification: <u>AI Module Profiles</u> (**MPAI-PRF**) enables an AIM instance to signal its Attributes – such as input data, output data, and functionality – and Sub-Attributes – such as languages supported by a Text and Speech Translation AIM – that uniquely characterise the AIM. Currently, MPAI-PRF defines the Attributes of eight AIMs but Profiles for more AIMs are likely to be defined in the future.

The effectiveness of the functions performed by an AIM can also be enabled or enhanced if the AIM knows more about the **characteristics of the Data** received. Examples of characteristics include:

- The CIE 1931 colour space of an instance of the Visual Data Type.
- The MP3 format of a speech segment.
- The WAV file format of an audio segment.
- The gamma correction applied to the device that produced a video.
- The Instance ID of an object in an audio segment.
- The Text conveyed by a speech segment.

Technical Specification: Data Types, Formats, and Attributes (MPAI-TFA) V1.2 specifies the Qualifier Data Type, a container that can be used to represent, for instance, that a Visual Data Type instance:

- Uses a given colour space (Sub-Type)
- Was produced by an AVC codec (Format).
- Is described by Dublin Core Metadata (Attribute).

Therefore, Qualifiers are a specialised type of metadata intended to support the operation of AIMs receiving data from other AIMs and conveying information on Sub-Types, Formats, and Attributes related to the Content. Qualifiers are intended to convey information for use by an AIM. They are human-readable but intended only to be used by AIMs. The combination of "Content" (the Data of a Data Type) and "Qualifier" (the combination of Sub-Type, Format, and Attributes) is called "Object".

MPAI provides a standard method to attach information to a Data Type instance called Annotation defined as Data attached to an Object or a Scene. As opposed to Qualifier that describes intrinsic properties of a Data Type, an Annotation is spatially and temporally local and changeable.

MPAI plans of publishing new versions of MPAI-TFA because of the large variety of applications requiring Qualifiers and the need for extending existing Qualifiers. MPAI-TFA users may communicate their need for extension of existing and specification of additional Data Type Qualifiers to the MPAI Secretariat. Therefore, versioning of Qualifiers is a critical component of MPAI-TFA.

The Chapters, Sections, and Annexes of this Technical Specification are Normative unless they are explicitly labelled as Informative. In all Chapters and Sections, Terms beginning with a capital letter are defined in *Table 1* if they are specific to this Technical Specification and in *Table 2* if they are common to all MPAI Technical Specifications. All Chapters and Annexes are Normative unless they are labelled as Informative.

3   Scope
Technical Specification: Data Types, Formats and Attributes (MPAI-TFA) V1.2 – in the following also called MPAI-TFA V1.2 or MPAI-TFA – specifies Qualifiers, i.e., additional information to an instance of a Data Type, a particular type of Data e.g., Text, Speech, and Visual that an AI Module may need to properly process the instance.
MPAI-TFA classifies that additional information as:
1. *Sub-Types*, information related to the different forms that can be taken by a Data Type instance, for example, Colour Space is a Sub-Type of the Visual Data Type.
2. *Formats*, the different ways in which a Data Type can be digitally represented or transported, for example, AAC is a Format of the Speech Data Type.

3. *Attributes*, the different types of information providing details on a Data Type instance, for example, the ID of an Object in a picture.

Sub-Types, Formats, and Attributes are further organised into subordinate hierarchies. New elements of the hierarchy may be added to this Technical Specification based on requests coming from application domains.

All information elements of a Qualifier are optional. The decision to add a particular element of the Qualifiers defined by this Technical Specification resides solely with the user.

The current version of MPAI-TFA specifies Qualifiers for the following Data Types: Text, Speech, Audio, and Visual.

This Technical Specification has been developed by CAE-DC, MMC-DC, PAF-DC, XRV-DC, and the CAV, MMM, and OSD groups of the Requirements Standing Committee. MPAI may publish new versions of MPAI-TFA or new standards covering – or extending – the scope currently covered by MPAI-TFA.

4 Definitions

Terms beginning with a <u>capital</u> letter have the meaning defined in *Table 1*. Terms applicable to all MPAI Technical Specifications are defined in *Table 2*.

Terms beginning with a <u>small</u> letter have the meaning commonly defined for the context in which they are used. For instance, *Table 1* defines *Object* and *Scene* but does not define *object* and *scene*.

A dash "-" preceding a Term in *Table 1* indicates the following readings according to the font:
1. Normal font: the Term in the table without a dash and preceding the one with a dash should be read <u>before</u> that Term. For example, "Avatar" and "- Model" will yield "Avatar Model."
2. *Italic* font: the Term in *Table 1* without a dash and preceding the one with a dash should be read <u>after</u> that Term. For example, "Avatar" and "- Portable" will yield "Portable Avatar."

Table 1 – Terms and Definitions relevant to MPAI-TFA

Terms	Definitions
Attribute	Information describing the features of a Data Type instance in addition to Sub-Type and Format.
Audio	A Data Type an instance of which represents analogue signals – or is rendered to be perceived – in the human-audible range (16 Hz – 20 kHz).
Data Type	A type of Data, such as Text, Speech, Audio, and Visual.
Dynamic	Refers to a Data Type that is time dependent.
Format	Information about the digital representation of a Sub-Type.
Machine Learning	A Process using training data to create a Model able to perform specific tasks such as classification.
Real	Refers to a Data Type instance that has been produced in a Real Space.
Qualifier	A Data Type including Sub-Types (e.g., colour space information), Format (e.g., compression and transport), and Attributes (e.g., semantic information) of a Data Type instance.
Space	
– *Real*	A space that is part of the Universe, i.e., the real world.
– *Virtual*	A space generated and maintained by a computing platform that can be rendered.

Speech	A Data Type an instance of which represents – or is rendered to be perceived – as an analogue signal with vocal characteristics.
Static	Refers to a Data Type that is not time dependent.
Sub-Type	A sub-category within a Data Type.
Synthetic	Refers to a Data Type instance that has been produced in a Virtual Space.
Text	A series of characters drawn from the finite alphabet of a Character Set.
Visual	A Data Type an instance of which represents analogue signals – or is rendered to be perceived – in the human-visible range (380 to 700 nanometres).
Universe	A synonym of the "real world".

Table 2 defines the capitalised Terms used in this standard that are not already included in Table 1.
NOTE: A hyphenated entry for e.g., "- Testing" should be read as adding that word to the closest non-hyphenated entry above it – in this case, "Conformance," giving "Conformance Testing" as the complete entry name.

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 Model (AIM)	A data processing element receiving AIM-specific Inputs and producing AIM-specific Outputs according to according to its Function. An AIM may be an aggregation of AIMs.
AI Workflow (AIW)	A structured aggregation of AIMs implementing a Use Case receiving AIW-specific inputs and producing AIW-specific outputs according to the AIW Function.
Application Standard	An MPAI Standard designed to enable a particular application domain.
Assessment Laboratory	A laboratory accredited to Assess the Grade of Performance of Implementations.
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.
Component	One of the 7 AIF elements: Access, Communication, Controller, Internal Storage, Global Storage, Store, and User Agent
Composite AIM	An AIM aggregating more than one AIM.
Component	One of the 7 AIF elements: Access, Communication, Controller, Internal Storage, Global Storage, Store, and User Agent
Conformance	The attribute of an Implementation of being a correct technical Implementation of a Technical Specification.

– Testing	The normative document specifying the Means to Test the Conformance of an Implementation.
– 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	Information in digital form.
– Format	The standard digital representation of Data.
– Type	An instance of Data with a specific Data Format.
– Semantics	The meaning of Data.
Descriptor	Coded representation of a text, audio, speech, or visual feature.
Digital Representation	Data corresponding to and representing a physical entity.
Ecosystem	The ensemble of actors making it possible for a User to execute an application composed of an AIF, one or more AIWs, each with one or more AIMs potentially sourced from independent implementers.
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.
AIM/AIW Storage	A Component to store data of the individual AIMs.
Identifier	A name that uniquely identifies an Implementation.
Implementation	1. An embodiment of the MPAI-AIF Technical Specification, or 2. An AIW or AIM of a particular Level (1-2-3) conforming with a Use Case of an MPAI Application Standard.
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 entity appointed by MPAI to assign ImplementerID's to Implementers.
Instance ID	Instance of a class of Objects and the Group of Objects the Instance belongs to.
Interoperability	The ability to functionally replace an AIM with another AIW having the same Interoperability Level
– Level	The attribute of an AIW and its AIMs to be executable in an AIF Implementation and to: 1. Be proprietary (Level 1) 2. Pass the Conformance Testing (Level 2) of an Application Standard 3. Pass the Performance Testing (Level 3) of an Application Standard.
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.
– Assessment	The normative document specifying the Means to Assess the Grade of Performance of an Implementation.
– Assessment Means	Procedures, tools, data sets and/or data set characteristics to Assess the Performance of an Implementation.
– Assessor	An entity Assessing the Performance of an Implementation.
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.
Scope	The domain of applicability of an MPAI Application Standard
Service Provider	An entrepreneur who offers an Implementation as a service (e.g., a recommendation service) to Users.
Standard	A set of Technical Specification, Reference Software, Conformance Testing, Performance Assessment, and Technical Report 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: 1.　The formats of the Input/Output data of the AIWs implementing the AIWs. 2.　The Connections of the AIMs of the AIW. 3.　The formats of the Input/Output data of the AIMs belonging to the AIW.
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 cybersecurity model primarily focused on data and service protection that assumes no implicit trust.

5　References

5.1　Normative References

1. Technical Specification: Governance of the MPAI Ecosystem (MPAI-GME) V1.1.
2. MPAI; Technical Specification; AI Framework (MPAI-AIF) V2.1.
3. MPAI; Technical Specification: Context-based Audio Enhancement (MPAI-CAE) – Use Cases (CAE-USC) V2.2.
4. MPAI; Technical Specification; Human and Machine Communication (MPAI-MMC) V1.1.
5. MPAI; Technical Specification; Multimodal Conversation (MPAI-MMC) V2.2.
6. MPAI; Technical Specification: MPAI Metaverse Model (MPAI-MMM) – Architecture (MMM-ARC) V1.2.
7. MPAI; Technical Specification: MPAI Metaverse Model (MPAI-MMM) – Technologies (MMM-ARC) V1.0.
8. MPAI; Technical Specification: Object and Scene Description (MPAI-OSD) V1.0.
9. MPAI; Technical Specification: AI Module Profiles (MPAI-PRF) V1.0.

5.2 Informative References

10. MPAI; Moving Picture, Audio, and Data Coding by Artificial Intelligence (MPAI)
11. MPAI; MPAI Standards.
12. MPAI; MPAI Patent Policy.

6   Data Types

MPAI-TFA V1.1 specifies Qualifiers for the following Data Types:

Media	Text	Speech	Audio	Visual	3D Model	Audio-Visual
Metaverse	Contract	Discovery	Information	Interpretation	Location	Program
	Rules					
Machine Learning	Model					