



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

Public document

N1731

2024/04/17

Source MMC-DC

Title Technical Specification: AI Module Profiles (MPAI-PRF) WD0.3 for Community
Comments

Target MPAI Members

This is WD 0.3 of Technical Specification: AI Module Profiles (MPAI-PRF) published for Community Comments. Anybody can submit comments to the draft by sending an email to the [MPAI secretariat](#) by 2024/05/08T23:59. MPAI will consider each comment received and may include some of them in the final version of MPAI-PRF.



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

MPAI Technical Specification

AI Module Profiles MPAI-PRF

WD0.3 for Community Comments

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 Annex 1 - Notices and Disclaimers.

MPAI Technical Specification

AI Module Profiles (MPAI-PRF) V1.0

(under development)

Contents

Foreword (Informative).....	4
1 Introduction (Informative).....	5
2 Scope	5
3 Terms and Definitions	5
4 References	7
4.1 Normative References	7
4.2 Informative References	8
5 AIMs requiring profiles.....	8
6 Profile signalling	8
6.1 Syntax.....	10
6.2 Semantics	10
7 AI Module Profiles.....	11
7.1 Entity Context Understanding (HMC-ECU).....	11
7.1.1 Definition	11
7.1.2 Specification.....	11
7.1.3 Attributes	12
7.2 Entity Dialogue Processing (MMC-EDP).....	12
7.2.1 Definition	12
7.2.2 Specification.....	Error! Bookmark not defined.
7.2.3 Attributes	13
7.3 Natural Language Understanding (MMC-NLU).....	13
7.3.1 Definition	13
7.3.2 Specification.....	14
7.3.3 Attributes	14
7.4 Personal Status Extraction (MMC-PSE).....	14
7.4.1 Definition	14
7.4.2 Specification.....	15
7.4.3 Attributes	15
7.5 Text and Speech Translation	15
7.5.1 Definition	16
7.5.2 Specification.....	16
7.5.3 Attributes	16
7.6 Audio-Visual Scene Rendering (PAF-AVR).....	17
7.6.1 Definition	17
7.6.2 Specification.....	17
7.6.3 Attributes	17
7.7 Personal Status Display (PAF-PSD)	18
7.7.1 Definition	18
7.7.2 Specification.....	18
7.7.3 Attributes	18
7.8 Text-to-Speech (MMC-TTS)	19

7.8.1	Definition	19
7.8.2	Specification.....	19
7.8.3	Attributes.....	19
Annex 1	- Notices and Disclaimers Concerning MPAI Standards (Informative).....	20
Annex 2	- General MPAI Terminology	22
Annex 3	- Patent Declarations.....	26

Foreword (Informative)

In recent years, Artificial Intelligence (AI) and related technologies have been applied to a broad range of applications, have started affecting the life of millions of people and are expected to do so even more in the future. As digital media standards have positively influenced industry and billions of people, so AI-based data coding standards are expected to have a similar positive impact. Moving Picture, Audio, and Data Coding by Artificial Intelligence (MPAI) has been established to develop standards that promote the efficient use of data especially using Artificial Intelligence technologies.

The accomplishment of the MPAI mission is facilitated by two foundational Technical Specifications:

- [Technical Specification: Artificial Intelligence Framework \(MPAI-AIF\)](#) [2] specifying an environment enabling initialisation, dynamic configuration, and control of AIWs in the standard AI Framework environment depicted in Figure 1 where AI applications called AI Workflows (AIW) can be executed. An AIW can include interconnected AI Modules (AIM).

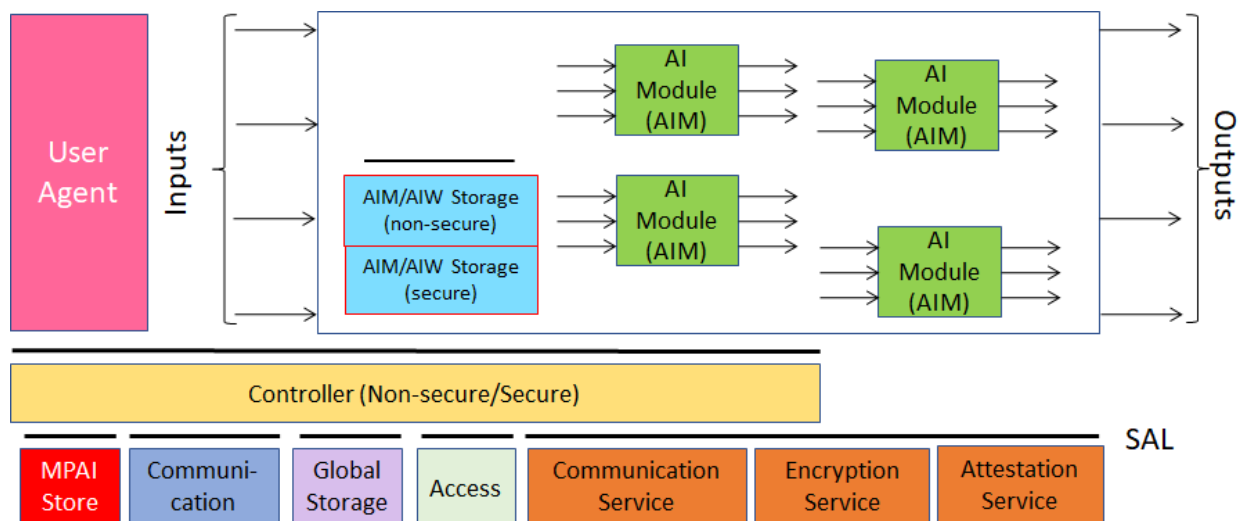


Figure 1 - The AI Framework (MPAI-AIF) V2 Reference Model

MPAI-AIF supports small- and large-scale high-performance components and promotes solution with improved explainability.

- [Technical Specification: Governance of the MPAI Ecosystem](#) [1] defining the operation of the following elements of the MPAI Ecosystem:
 - Standards, i.e., the ensemble of Technical Specifications, Reference Software, Conformance Testing, and Performance Assessment.
 - Implementers of MPAI Technical Specifications.
 - MPAI Store in charge of making AIMs and AIWs submitted by Implementers available to Integrators and End-Users.

- Performance Assessors, independent entities assessing the performance of implementations in terms of Reliability, Replicability, Robustness, and Fairness.
- End Users.

1 Introduction (Informative)

Some AI Modules (AIM) receive more/less input and produce more/fewer output data than an AIM with the same name that is used in other AI Workflows (AIW) even though they nominally perform the same functions. For instance, the [Natural Language Understanding](#) (MMC-NLU) AIM in the [Conversation with Emotion](#) AIW (MMC-CWE) receives [Recognised Text](#) as input and produces [Refined Text](#) and [Meaning](#). However, the MMC-NLU AIM used in the [Communicating Entities in Context](#) AIW (HMC-CEC) may also use [Audio-Visual Scene Geometry](#) and Audio and/or Visual Object [Instance IDs](#) to achieve better understanding of the textual sentence.

Since it is not realistic to require that all MMC-NLU AIMs be equipped with the additional logic required to exploit Geometry and Object information to provide Refined Text and Meaning, *Technical Specification: AI Module Profiles (MPAI-PRF)* provides a mechanism that unambiguously signals which characteristics of an AIM – called Attributes in the following – are supported by the AIM.

Table 1 defines the Terms specific of this Technical Specification and Table 5 those used across MPAI Technical Specifications. Chapters, Sections, and Annexes are Normative unless they are explicitly identified as Informative.

2 Scope

Technical Specification: AI Module Profiles (MPAI-PRF) enables the signalling of AI Module Attributes – input data, output data or functionality – that uniquely characterise an AIM instance. An [AIM Profile](#) is a label that uniquely identifies the set of AIM Attributes that are either supported or not supported by an AIM instance.

MPAI-PRF has been developed by the Multimodal Conversation (MPAI-MMC) and Portable Avatar Format (MPAI-PAF) Development Committees. MPAI may issue new Versions of this Technical Specification or new Technical Specifications.

3 Definitions

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 before that Term. For example, “Data” and “- Type” means “Avatar Model.”
2. *Italic* font: the Term in the table without a dash and preceding the one with a dash should be read after that Term. For example, “AI Module” and “- *Basic*” means “Basic AI Module.”

The definition of all MPAI Terms is found [online](#).

Table 1 - Terms used in this Technical Specification

Term	Definition
AI Module (AIM)	A data processing component performing a Function by processing AIM-specific Input Data and producing AIM-specific Output Data.

- Profile	The label that uniquely identifies a set of Attributes.
- Attribute	An input Data or an output Data or a functionality, such as the ability to translate.
- <i>Basic</i>	An AIM that does not aggregate other AIMs.
- <i>Composite</i>	An AIM aggregating more than one AIM.
Avatar	An Object rendered to represent a Human or a Machine in a virtual space.
- Model	An inanimate Avatar exposing animation interfaces.
- <i>Portable</i>	A Data Type including Avatar ID, Time, Visual Environment, Spatial Attitude, Avatar Model, Body Descriptors, Face Descriptors, Language Preference, Speech Coding, Speech Data, Text, and Personal Status.
Body	A digital representation of a human body.
- Object	A Data Type representing the body of an Entity, head included, face excluded.
- Descriptors	A Data Type representing the features of an Entity's Body.
Context	Information surrounding an Entity and providing additional insight into the information the Entity communicates.
Data	Information in digital form.
- Format	A standard representation of Data.
- Type	An instance of Data with a specific Data Format
Entity	A human in a real environment or digitally represented as a Digitised Human in a Virtual Environment a Digital or a Virtual Human in a Virtual Environment.
Face	A digital representation of a human face.
- Descriptors	A Data Type representing the motion and conveying information on the Personal Status of the face of a human or an avatar.
- Objects	A Data Type representing the face of an Entity.
Factor	One of Cognitive State, Emotion, and Social Attitude
Modality	One of Text, Speech, Face, or Gesture.
Object	Data that can be rendered to cause an Experience.
- <i>Audio</i>	A Data Type representing an object or a computer-generated Object that can be rendered to and perceived by a human ear.
- <i>Audio-Visual</i>	An Object composed of Audio and Visual Objects sharing the same Spatial Attitude.
- Instance	The instance of an Audio Object.
- <i>Visual</i>	The digital representation of an object captured by an electromagnetic or high-frequency audio signal or computer-generated that can be rendered to and perceived by a human eye.
Personal Status	A Data Type representing the ensemble of information internal to a person expressed by 3 Factors (Cognitive State, Emotion, Social Attitude) conveyed by one or more Modalities (Text, Speech, Face, and Gesture).
Point of View	The Spatial Attitude of an Entity user looking at an Environment.
Scene Descriptors	The digital representation of the features of a scene.
<i>Audio</i>	A Data Type representing the Audio Objects and their spatial arrangement in an Audio Scene.
<i>Audio-Visual</i>	A Data Type representing the Audio-Visual Objects and their spatial arrangement in an Audio-Visual Scene.
<i>Visual</i>	A Data Type representing the Visual Objects and their spatial arrangement in a Visual Scene.

Scene Geometry	The digital representation of the Object arrangement of a Scene.
- <i>Audio</i>	A Data Type representing the spatial arrangement of the Audio Objects of a Scene.
- <i>Audio-Visual</i>	A Data Type representing the spatial arrangement of the Audio, Visual, and Audio-Visual Objects of a Scene.
- <i>Visual</i>	A Data Type representing the spatial arrangement of the Visual Objects of a Scene.
Speech	Digital representation of analogue speech sampled at a frequency between 8 kHz and 96 kHz with a number of bits/sample of 8, 16 or 24, and non-linear and linear quantisation or compressed. Data with characteristics of Speech may be synthetically produced.
- Descriptors	A Data Type representing information elements incorporated in a Speech Segment, e.g., personal identity, Personal Status, additional factors such as vocal tension, creakiness, whispery quality, etc.
- Model	A Neural Network trained to generate utterances with specific Speech Descriptors.
- Object	An Object described by Speech Descriptors.
Text	A series of characters drawn from a finite alphabet of a Character Set.
- Descriptors	A Data Type including the digital representation of the features of Text.
- Object	A string of Text.
- <i>Recognised</i>	The Text produced by the Automatic Speech Recognition AIM.

4 References

4.1 Normative References

- 1 MPAI; Technical Specification; MPAI Ecosystem Governance (MPAI-GME) V1.1; <https://mpai.community/standards/mpai-gme/>.
- 2 MPAI; Technical Specification: AI Framework (MPAI-AIF) V2.0; <https://mpai.community/standards/mpai-aif/>.
- 3 MPAI; Technical Specification: Human and Machine Communication (MPAI-HMC) V1.0; <https://mpai.community/standards/mpai-hmc/>.
 - 3.1 Entity and Context Understanding: <https://mpai.community/standards/aiws-and-aims/entity-context-understanding-hmc-ecu/>.
- 4 MPAI; Technical Specification: Multimodal Conversation (MPAI-MMC) V2.1; <https://mpai.community/standards/mpai-mmc/>.
 - 4.1 Entity Dialogue Processing: <https://mpai.community/standards/aiws-and-aims/entity-dialogue-processing-hmc-edp/>.
 - 4.2 Natural Language Understanding: <https://mpai.community/standards/aiws-and-aims/natural-language-understanding-hmc-nlu/>.
 - 4.3 Personal Stats Extraction: <https://mpai.community/standards/aiws-and-aims/personal-status-extraction-mmc-pse/>.
 - 4.4 Text and Speech Translation: <https://mpai.community/standards/aiws-and-aims/text-and-speech-translation-mmc-tst/>.
 - 4.5 Text To Speech: MMC-TTS is specified at <https://mpai.community/standards/aiws-and-aims/text-to-speech-mmc-tts/>.
- 5 MPAI; Technical Specification: Portable Avatar Format (MPAI-PAF) V1.1; <https://mpai.community/standards/mpai-paf/>.

- 5.1 Audio-Visual Rendering: <https://mpai.community/standards/aiws-and-aims/audio-visual-scene-rendering-hmc-avr/>.
- 5.2 Personal Status Display: PAF-PSD is specified at <https://mpai.community/standards/aiws-and-aims/personal-status-display-paf-psd/>.
- 6 ISO 639:2023 - Code for individual languages and language groups; <https://www.iso.org/standard/74575.html>

4.2 Informative References

- 7 MPAI; The MPAI Statutes; <https://mpai.community/statutes/>.
- 8 MPAI; The MPAI Patent Policy; <https://mpai.community/about/the-mpai-patent-policy/>.

5 AIMs requiring profiles

The first column of Table 2 provides the current list of names of AIMs requiring Profiles, the second column their acronyms, and the third column the JSON Metadata [2].

Table 2 - AIMs requiring Profiles

Name	AIMs	JSON
Audio-Visual Scene Rendering	PAF-AVR	X
Entity Context Understanding	HMC-ECU	X
Entity Dialogue Processing	MMC-EDP	X
Natural Language Understanding	MMC-NLU	X
Personal Status Display	PAF-PSD	X
Personal Status Extraction	MMC-PSE	X
Text and Speech Translation	MMC-TST	X
Text-to-Speech	MMC-TTS	X

This list may be subject to future revision because:

1. Some AIMs currently having a single form, may subsequently found to require Profiles.
2. The functionality of existing AIMs may be extended in a way that requires Profiles.
3. New AIMs may be specified that require Profiles.

Note that this document does not distinguish between Composite and Basic AIMs.

6 Profile signaling

The 23 Attributes of Table 3 represented with 3-character codes have been found necessary to identify Attributes of the AI Modules of Table 2.

Table 3 - Coding of AIM Attributes

Attributes	Code
Audio Object	AUO
Audio-Visual Scene Geometry	AVG
Avatar Model	AVM
Body Object	BDO
Body Descriptors	BDD
Face ID	FCI

Face Object	FCO
Face Descriptors	FCD
Input Personal Status	IPS
Language Preferences	LGP
Memory	MEM
Speech Model	SPM
Object Instance ID	OII
Point of View	POV
Portable Avatar	PAV
Speech Descriptors	SPD
Speaker ID	SPI
Speech Object	SPO
Text Descriptors	TXD
Text Object	TXO
Recognised Text	TXR
Translation	TRN
Visual Object	VIO

Thus, a specific AIM Profile will require:

1. The three characters identifying the Technical Specification that specifies the AIM.
2. The three characters identifying the AIM of that Technical Specification.
3. The Version and Subversion of the Technical Specification.
4. The Profile-specific sequence of coded Attributes drawn from Table 3.

For instance, the Profile of a Natural Language Understanding (HMC-NLU) AIM that does not handle spatial information (see Section 8.3) is labelled in two ways, allows more compact signalling matched to the number of Attributes supported by an AIM:

Removing unsupported Attributes MMC-NLU-V2.1(ALL-AVG-OII)
Adding supported Attributes MMC-NLU-V2.1(NUL+TXO+TXR)

Attributes, however, are not always sufficient to identify the capabilities of an AIM instance. For instance, an AIM instance of Personal Status Display (PAF-PSD) may support Personal Status, but only the Speech (PS-Speech) and Face (PS-Face) Personal Status Factors. This is illustrated by the following two examples:

Removing unsupported Attributes PAF-PSD-V1.1(ALL@IPS#SPE#FCE)
Adding supported Attributes PAF-PSD-V1.1(NUL+TXT+AVM@IPS#FCE#GST)

Here @ prefixed to IPS signals that the AIM supports Personal Status, but only of Speech and Face and of Face and Gesture represented by PSS, PSF, and GST, the codes of the PS-Speech, PS-Face, and PS-Gesture Sub-Attributes, respectively (the full list of Personal Status Sub-Attributes is provided by Table 4). The second case may apply for a sign-language capable AIM.

Currently, MPAI-PRF supports another use of Sub-Attributes. The Test and Speech Translation AIM needs Sub-Attributes to signal which languages are supported in which direction as exemplified below:

Removing unsupported Attributes MMC-TST-V2.1(NUL@TRN#eng→ita)
Adding supported Attributes MMC-TST-V2.1(ALL-ISD@TRN#kor↔fra#ger→swa)

The first case refers to an AIM that only supports text translation from English to Italian and the second to an AIM that does not support Speech Descriptors but supports text and speech translation from both Korean to and from French, and from German to Swahili (the TRN Sub-Attributes are specified in Table 4).

The capabilities of an AIM can thus be represented along two dimensions: the first relates to its Attributes and is called Profile and the second relates to its Sub-Attributes and is called Level. Table 4 lists the Sub-Attributes of the Personal Status and Translation Attributes:

Table 4 - Attributes and Sub-Attributes

Attribute	Coded Attributes	Sub-Attributes	Coded Sub-Attributes
Personal Status	IPS	Text	TXT
		Speech	SPE
		Face	FCE
		Gesture	GST
Translation	TRN	Language	3-letter codes of [5.1], Part 3

An AIM implementing Attribute signalling does not expose the missing Attribute interfaces.

7 JSON Syntax and Semantics

The following syntax and semantics can be used to define a file that may contain many Attributes and Sub-Attributes.

7.1 Syntax

```
{
  "$schema": "http://json-schema.org/draft/2019-09/schema",
  "title": "AIMProfiles",
  "type": "object",
  "properties": {
    "Header": {
      "type": "string",
      "pattern": "^PRF-[0-9]{1,2}[.][0-9]{1,2}$"
    },
    "Profile": {
      "type": "string",
      "pattern":
        "^(ALL|NUL)([+-](AUO|AVG|AVM|BDO|BDD|FCI|FCO|FCD|IPS|LGP|MEM|SPM|OII|POV|PAV|SPD|SPI|SPO|TXD|TXO|
        TXR|TRN|VIO)|@IPS(#[TXT|SPE|FCE|GST])+|@TRN(#[a-z]{3})(<->|->)([a-z]{3}))+$"
    }
  }
}

{
  "Header": "PRF-1.0",
  "Profile": "NUL+TXO@IPS#SPE#FCE"
}
```

7.2 Semantics

Label	Size	Description
HEADER	9 Bytes	
• Standard	3 Bytes	The three characters “PRF”
• Dash-separator	1 Byte	The character “-”
• Version	1 Byte	Major version
• Dot-separator	1 Byte	The character “.”

• Subversion	1 Byte	Minor version
Profile[]	N*4 Bytes	N is the number of Attributes
• Dash-separator	1 Byte	The character “-”
• MissingAttribute[]	3 Bytes	An Attribute of Table 3
○ Plus separator	1 Byte	The character “+”
○ AttributeCode		
• “AddedAttribute[]	3 Bytes	An Attribute of Table 3
○ Hash-separator	1 Bytes	The character “#”
○ AttributeCode	3 Byte	Sub-Attributes of Table 3

8 AIM Profiles

8.1 Entity Context Understanding (HMC-ECU)

8.1.1 Definition

HMC-ECU

- 1 Receives Audio-Visual Scene Descriptors.
- 2 Processes the Descriptors to enable Machine to achieve understanding of the information conveyed by an Entity and its Context.
- 3 Produces
 - 3.1 Personal Status
 - 3.2 Refined and Translated Text
 - 3.3 Meaning
 - 3.4 An Audio Instance ID
 - 3.5 A Visual Instance ID
 - 3.6 The Geometry of the Scene that contains the Audio and Visual Objects.
- 4 Enables the downstream Entity Dialogue Processing AIM to produce a pertinent Communication Item as a Portable Avatar.

8.1.2 Specification

Entity and Context Understanding is specified at [3.1].

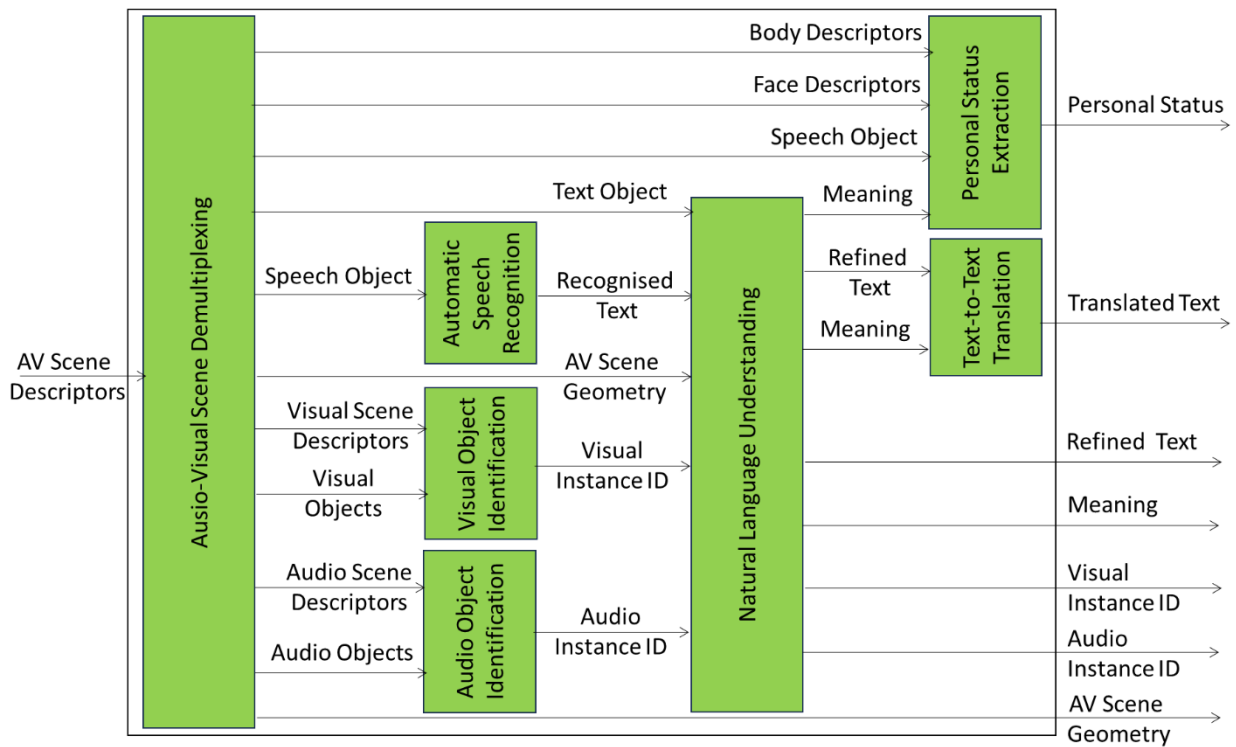


Figure 2 - Entity and Context Understanding

8.1.3 Attributes

An ECU AIM Profile is determined by whether the AIM uses one or more of the following attributes:

Attribute	Code	Function
Body Descriptors	BDD	HMC-ECU Receives Body Descriptors
Face Descriptors	FCD	HMC-ECU Receives Face Descriptors
Speech Object	SPO	HMC-ECU Receives Speech Object
Text Object	TXO	HMC-ECU Receives Text Object
Visual Object	VIO	HMC-ECU Receives Visual Object
Audio Object	AUO	HMC-ECU Receives Audio Object
Audio-Visual Scene Descriptors	AVS	HMC-ECU Receives Audio-Visual Scene Descriptors
Audio-Visual Scene Geometry	AVG	HMC-ECU Receives Audio-Visual Scene Geometry
Translation	TRN	HMC-ECU Translates Text Object

8.2 Entity Dialogue Processing (MMC-EDP)

8.2.1 Definition

MMC-EDP

- 1 Receives:
 - 1.1 Text Object
 - 1.2 Object Instance ID
 - 1.3 Input Personal Status
 - 1.4 Text Descriptors
 - 1.5 AV Scene Geometry
 - 1.6 Speaker ID
 - 1.7 Face ID

- 1.8 Memory
- 2 Processes the information received
 - 2.1 Handling one Speech Object at a time.
 - 2.2 Taking past Speech Objects into account.
- 3 Produces elements of the Machine Response to the data issued by the Entity in its Context in the form of:
 - 3.1 Text.
 - 3.2 Personal Status

8.2.2 Specification

MMC-EDP is specified at [4.1].

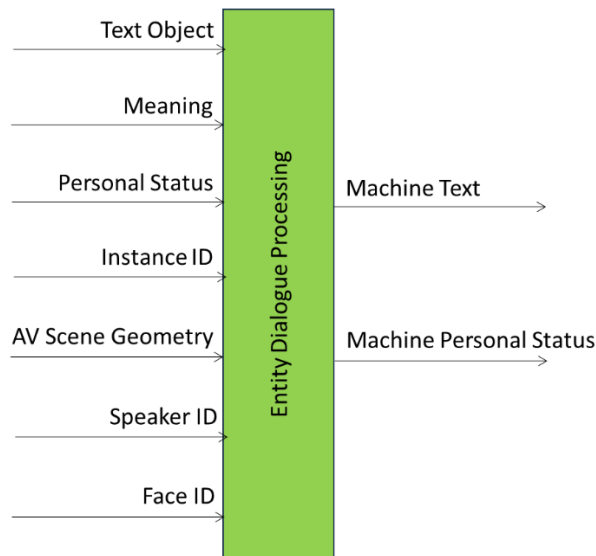


Figure 3 - Entity Dialogue Processing

8.2.3 Attributes

A Profile is determined by whether the AIM uses one or more of the following attributes:

Attribute	Code	Function
Text Object	TXO	MMC-EDP receives Text (directly from human or through NLU).
Object Instance ID	OII	MMC-EDP receives the ID of an A/V/AV Instance referenced in the dialogue.
Input Personal Status	IPS	MMC-EDP receives Personal Status.
Text Descriptors	TXD	MMC-EDP receives Meaning.
AV Scene Geometry	AVG	MMC-EDP receives AV Scene Geometry to enable it to locate the Object.
Speaker ID	SPI	MMC-EDP receives Speaker ID.
Face ID	FCI	MMC-EDP receives Face ID.
Memory	MEM	MMC-EDP takes into account prior Input Data of the dialogue session.

8.3 Natural Language Understanding (MMC-NLU)

8.3.1 Definition

MMC-NLU

- 1 Receives
 - 1.1 Text Object directly input by the Entity.
 - 1.2 Recognised Text from the Automatic Speech Recognition AIM.
 - 1.3 An ID of an Instance.
 - 1.4 The Audio-Visual Scene Geometry containing the Instance.
- 2 Refines the Input Text if coming from an Automatic Speech Recognition AIM and extracts the Meaning (Text Descriptors) from the Recognised Text or from a Text Object provided by the Entity.
- 3 Produces
 - 3.1 Refined Text.
 - 3.2 Text Descriptors (Meaning).
- 4 Enables the Personal Stats Display to produce a Portable Avatar.

8.3.2 Specification

MMC-NLU is specified at [4.2].

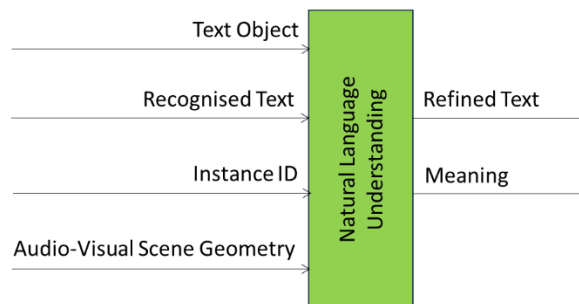


Figure 4 - Natural Language Understanding

8.3.3 Attributes

A Profile is determined by whether the AIM uses one or more of the following attributes:

Attribute	Code	Function
Text Object	TXO	MMC-NLU receives Text directly from human.
Recognised Text	TXR	MMC-NLU receives text from ASR.
Object Instance ID	OII	MMC-NLU receives Object Instance ID
Audio-Visual Scene Geometry	AVG	MMC-NLU receives AV Geometry.
Text Descriptors	TXD	MMC-NLU Produces Text Descriptors (Meaning)

8.4 Personal Status Extraction (MMC-PSE)

8.4.1 Definition

MMC-PSE

- 1 Receives
 - 1.1 Text information
 - 1.1.1 Text Selector informs about availability of Text Descriptors
 - 1.1.2 Text Object
 - 1.1.3 Text Descriptors
 - 1.2 Speech information
 - 1.2.1 Speech Selector informs about availability of Speech Descriptors
 - 1.2.2 Speech Object
 - 1.2.3 Speech Descriptors
 - 1.3 Face information informs about availability of Face Descriptors

- 1.3.1 Face Selector
- 1.3.2 Face Object
- 1.3.3 Face Descriptors
- 1.4 Body information
 - 1.4.1 Gesture Selector informs about availability of Gesture Descriptors
 - 1.4.2 Body Object
 - 1.4.3 Gesture Descriptors
- 2 Processes the received information
 - 2.1 Computing the Modality (Text, Speech, Face, and Gesture) Descriptors for Cognitive State, Emotion and Social Attitude if Modality Selector signals that it is not already available.
 - 2.2 Interpreting the Descriptors to produce the Personal Statuses of the Modalities.
 - 2.3 Multiplexing the Personal Statuses of the Modalities into the Personal Status.
- 3 Produces Personal Status.
- 4 Enables the Entity Dialogue Processing to improve its ability to respond.

8.4.2 Specification

MMC-PSE is specified at [4.3].

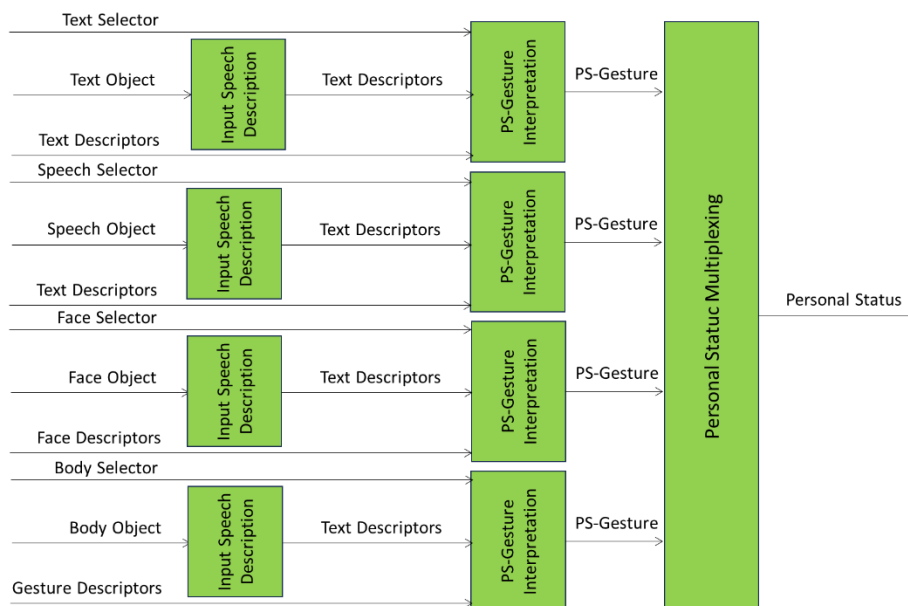


Figure 5 - Personal Status Extraction

8.4.3 Attributes

A PSE AIM Profile is determined by whether the AIM uses one or more of the following Attributes:

Attribute	Code	Function
Text Object	TXO	MMC-PSE receives Text
Speech Object	SPO	MMC-PSE receives Speech
Face Object	FCO	MMC-PSE receives Face
Body Object	BDO	MMC-PSE receives Gesture

When an MMC-PSE is used as a component AIM in a Composite AIM as in the case of HMC-ECU, the MMC-PSE Attributes become Sub-Attributes of the Composite AIM.

8.5 Text and Speech Translation (MMC-TST)

8.5.1 Definition

MMC-TST

- 1 Receives:
 - 1.1 Selector to inform whether
 - 1.1.1 The AIM output should be Text or Speech.
 - 1.1.2 The output Speech should retain the Features of the input Speech.
 - 1.2 Language Preferences in the form of requested input and output language.
 - 1.3 Personal Status.
 - 1.4 Text.
 - 1.5 Speech.
- 2 Performs (a subset of) the following:
 - 2.1 Converts input Speech into Text using Personal Status.
 - 2.2 Translates the Text to the target language
 - 2.3 Extracts the Features from Speech.
 - 2.4 Converts Text into Speech adding the Speech Features.
- 3 Produces:
 - 3.1 Translated Text
 - 3.2 Translated Speech

8.5.2 Specification

MMC-TST is specified at [4.4].

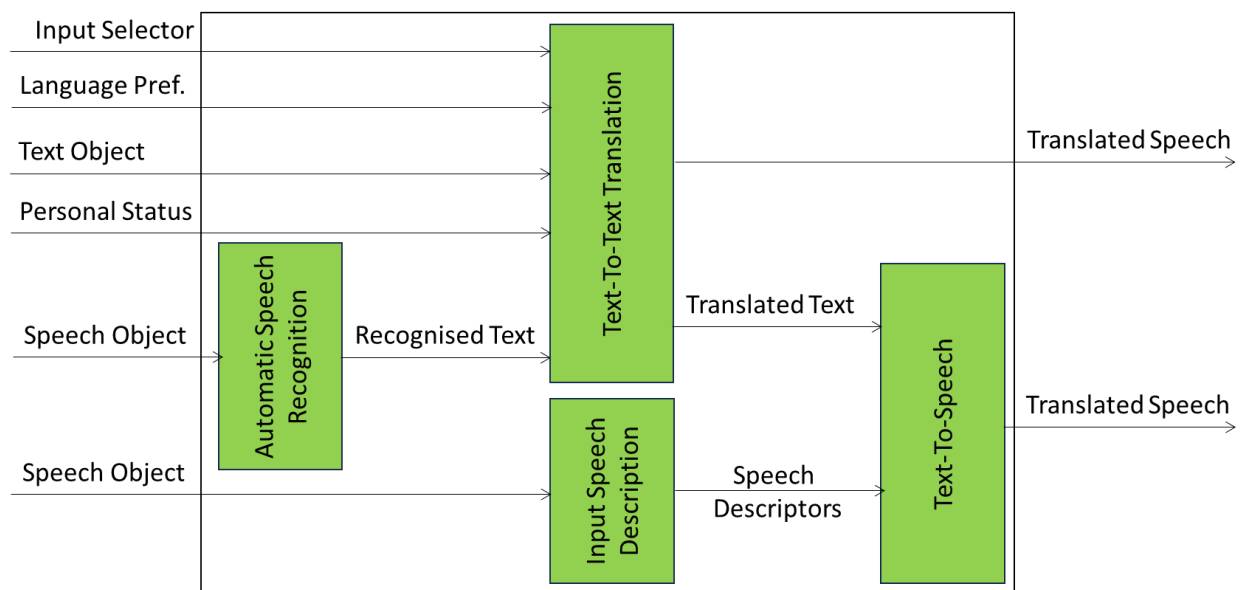


Figure 6 - Text and Speech Translation

8.5.3 Attributes

An MMC-TST AIM Profile is determined by whether the AIM uses one or more of the following Attributes:

Attributes	Code	Functions
Language Preferences	LGP	MMC-TST receives information on input and output languages.
Text Object	TXO	MMC-TST receives Text

Speech Object	SPO	MMC-TST receives Speech
Speech Descriptors	SPD	MMC-TST uses Speech Descriptors
Personal Status	IPS	MMC-TST receives Personal Status

When an MMC-TST is used as a component AIM in a Composite AIM as in the case of HMC-ECU, the LGP (Language Preferences) Attribute of MMC-TST become Sub-Attributes of the Composite AIM represented as 3-letter codes of [5.1], Part 3.

8.6 Audio-Visual Scene Rendering (PAF-AVR)

8.6.1 Definition

PAF-AVR

- 1 Receives
 - 1.1 Audio-Visual Scene Descriptors or a Portable Avatar.
 - 1.2 A Point of View.
- 2 Transforms the Portable Avatar into Audio-Visual Scene Descriptors.
- 3 Produces
 - 3.1 Text included in the Portable Avatar.
 - 3.2 Output Audio, the result of rendering the Audio Scene Descriptors from the Point of View.
 - 3.3 Output Visual, the result of rendering the Visual Scene Descriptors from the Point of View.

8.6.2 Specification

PAF-AVR is specified at [5.1].

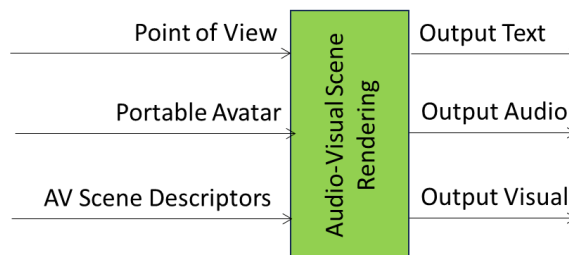


Figure 7 - Audio-Visual Scene Rendering

8.6.3 Attributes

A PAF-AVR AIM Profile is determined by whether the AIM uses one or more of the following attributes:

Attribute	Code	Function
Point of View	POV	PAF-AVR is informed to provide Output Audio and/or Output Visual as perceived from a Point of View.
Portable Avatar	PAV	PAF-AVR receives a Portable Avatar and produces an Audio-Visual Scene from the Point of View.
Audio-Visual Scene Descriptors	AVS	PAF-AVR receives Audio-Visual Scene Descriptors and produces an Audio-Visual Scene from the Point of View.
Output Text	TXO	PAF-AVR produces Text Object.
Output Audio	AUO	PAF-AVR produces Audio Object.
Output Visual	VIO	PAF-AVR produces Visual Object.

8.7 Personal Status Display (PAF-PSD)

8.7.1 Definition

PAF-PSD

- 1 Receives
 - 1.1 Text Object
 - 1.2 Personal Status
 - 1.3 Avatar Model
 - 1.4 Speech Model
 - 1.5 NN Format
- 2 Uses
 - 2.1 Text and PS-Speech to produce the Machine Speech.
 - 2.2 Machine Speech, Avatar Model, and PS-Face to produce Machine Face Descriptors.
 - 2.3 Machine Text, Avatar Model, and PS-Gesture to produce Machine Body Descriptors
- 3 Produces Portable Avatar.
- 4 Enables PAF-AVR to render the Portable Avatar produced by PAF-PSD.

8.7.2 Specification

PAF-PSD is specified at [5.2].

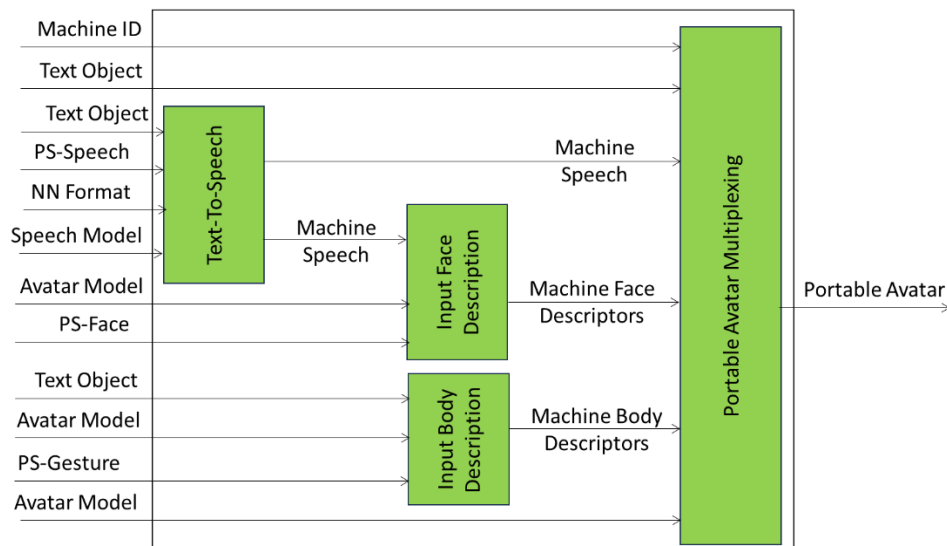


Figure 8 - Personal Status Display

8.7.3 Attributes

A PAF-PSD AIM Profile is determined by whether the AIM uses one or more of the following Attributes:

Attribute	Code	Function
Text Object	TXO	PAF-PSD receives Text and produces Speech.
Personal Status	IPS	PAF-PSD receives Personal Status.
Speech Model	SPM	
Avatar Model	AVM	PAF-PSD receives an Avatar Model.

8.8 Text-to-Speech (MMC-TTS)

8.8.1 Definition

MMC-TTS

- 1 Receives
 - 1.1 Text Object.
 - 1.2 Personal Status.
 - 1.3 Speech Model.
- 2 Feeds Text Object and Personal Status to Speech Model.
- 3 Produces an utterance.

8.8.2 Specification

MMC-TTS is specified at [4.5].

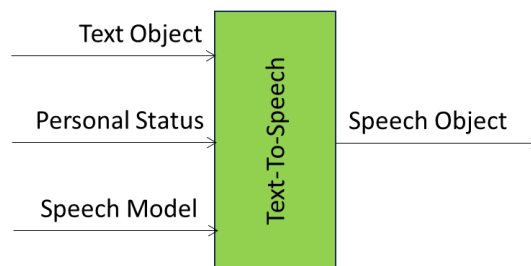


Figure 9 - Text-To-Speech

8.8.3 Attributes

An MMC-TTS AIM Profile is determined by whether the AIM uses one or more of the following Attributes:

Attribute	Code	Function
Text Object	TXO	MMC-TTS receives Text Object
Personal Status	IPS	MMC-TTS receives Personal Status
Speech Model	SPM	MMC-TTS receives NN Speech Model

Annex 1 - Notices and Disclaimers Concerning MPAI Standards (Informative)

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.

Copyright

MPAI draft and approved standards, whether they are in the form of documents or as web pages or otherwise, are copyrighted by MPAI under Swiss and international copyright laws. MPAI Standards are made available and may be used for a wide variety of public and private uses, e.g., implementation, use and reference, in laws and regulations and standardisation. By making these documents available for these and other uses, however, MPAI does not waive any rights in copyright to its Standards. For inquiries regarding the copyright of MPAI standards, please contact the [MPAI Secretariat](#).

The Reference Software of an MPAI Standard is released with the [MPAI Modified Berkeley Software Distribution licence](#). However, implementers should be aware that the Reference Software of an MPAI Standard may reference some third party software that may have a different licence.

Annex 2 - General MPAI Terminology

The capitalised Terms used in this standard that are not already included in Table 1 are defined in Table 5.

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 5 – Terms used across several MPAI Technical Specifications

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 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 Dataset	A dataset used to Test the Conformance of an implementation to a Technical Specification.
- Testing Means	Procedures, tools, data sets and/or data set characteristics to Test the Conformance of an Implementation.

- Testing Procedure	The sequence of steps to be performed to Test the Conformance of an implementation.
- Testing Tools	Devices and/or software used 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:

	<ol style="list-style-type: none"> 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 their 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	<p>(Framework) the normative specification of the AIF.</p> <p>(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:</p> <ol style="list-style-type: none"> 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.

Annex 3 - Patent Declarations

Technical Specification: AI Module Profiles has been developed using four MPAI Technical Specifications: MPAI-CAE, MPAI-HMC, MPAI-MMC, MPAI-OSD, and MPAI-PAF. While the AIM Profile identification method has been originally developed by MPAI and is freely available, the data types referenced by the MPAI-PRF Technical Specification may be affected by the Patent Coverage of the relevant Technical Specifications.