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

**Public document**

|  |  |
| --- | --- |
| **N1102** | 2023/02/22 |
| **Source** | Requirements (XRV) |
| **Title** | XR Venue (MPAI-XRV) Use Case and Functional Requirements WD0.9 |
| **Target** | MPAI-29 |

# Abstract

XR Venues (MPAI-XRV) is an MPAI project addressing a multiplicity of use cases enabled by Extended Reality (XR), the combination of Augmented Reality (AR), Virtual Reality (VR) and Mixed Reality (MR) technologies and enhanced by Artificial Intelligence (AI) technologies. The word venue is used as a synonym for real and virtual environments.

The goals of the project are:

1. To identify and describe XR Venue use cases.
2. Characterise components called AI Modules (AIMs) reusable across use cases.
3. To develop requirements for the workflows called AI Workflows (AIWs) implementing the identified use cases and for AIM functions and input/output data.
4. To present intermediate findings to the industry.
5. To draft and publish a Call for Technologies satisfying the identified functional requirements and commercial requirements.
6. To specify the XRV enabling technologies in a standard.

Contents

[Abstract 1](#_Toc126691664)

[1 Introduction 1](#_Toc126691665)

[2 A Real-Virtual Interaction Model 1](#_Toc126691666)

[3 Terms and definitions 1](#_Toc126691667)

[4 Use Cases 1](#_Toc126691668)

[4.1 Important Information 1](#_Toc126691669)

[4.2 eSports Tournament (XRV-EST). 1](#_Toc126691670)

[4.2.1 Purpose 1](#_Toc126691671)

[4.2.2 Description 1](#_Toc126691672)

[4.3 Live theatrical stage performance. 1](#_Toc126691673)

[4.3.1 Purpose 1](#_Toc126691674)

[4.3.2 Description and flow of actions 1](#_Toc126691675)

[4.4 Experiential retail/shopping. 1](#_Toc126691676)

[4.4.1 Purpose 1](#_Toc126691677)

[4.4.2 Description and flow of actions 1](#_Toc126691678)

[4.5 Collaborative immersive laboratory 1](#_Toc126691679)

[4.5.1 Purpose 1](#_Toc126691680)

[4.5.2 Description 1](#_Toc126691681)

[4.5.3 Microscopic dataset visualisation 1](#_Toc126691682)

[4.5.4 Macroscopic dataset visualisation and simulation 1](#_Toc126691683)

[4.5.5 Educational lab 1](#_Toc126691684)

[4.5.6 Collaborative CAD 1](#_Toc126691685)

[4.6 Immersive art experience. 1](#_Toc126691686)

[4.6.1 Purpose 1](#_Toc126691687)

[4.6.2 Description 1](#_Toc126691688)

[4.7 DJ/VJ performance at a dance party. 1](#_Toc126691689)

[4.7.1 Purpose 1](#_Toc126691690)

[4.7.2 Description 1](#_Toc126691691)

[4.8 Live concert performance. 1](#_Toc126691692)

[4.8.1 Purpose 1](#_Toc126691693)

[4.8.2 Description 1](#_Toc126691694)

[4.9 Experiential marketing/branding. 1](#_Toc126691695)

[4.9.1 Purpose 1](#_Toc126691696)

[4.9.2 Description 1](#_Toc126691697)

[4.10 Meetings/presentations. 1](#_Toc126691698)

[4.10.1 Purpose 1](#_Toc126691699)

[4.10.2 Description 1](#_Toc126691700)

[5 Functional requirements 1](#_Toc126691701)

[5.1 eSports Tournament (XRV-EST). 1](#_Toc126691702)

[5.1.1 Processing 1](#_Toc126691703)

[5.1.2 Operations 1](#_Toc126691704)

[5.1.3 Reference architecture 1](#_Toc126691705)

[5.1.4 AI Modules 1](#_Toc126691706)

[5.1.5 AIM I/O Data Formats 1](#_Toc126691707)

[5.2 Live theatrical stage performance 1](#_Toc126691708)

[5.2.1 Operation 1](#_Toc126691709)

[5.2.2 Reference Architecture 1](#_Toc126691710)

[5.2.3 AI Modules 1](#_Toc126691711)

[5.2.4 AIM I/O Data Formats 1](#_Toc126691712)

[5.3 Experiential retail/shopping 1](#_Toc126691713)

[5.3.1 Reference Architecture 1](#_Toc126691714)

[5.3.2 AI Modules 1](#_Toc126691715)

[5.3.3 AIM I/O Data Formats 1](#_Toc126691716)

[6 References 1](#_Toc126691717)

[6.1 Normative References 1](#_Toc126691718)

[6.2 Informative References 1](#_Toc126691719)

[Annex A Basics about MPAI 1](#_Toc126691720)

[Annex B MPAI-wide terms and definitions 1](#_Toc126691721)

[Annex C Notices and Disclaimers Concerning MPAI Standards (Informative) 1](#_Toc126691722)

[Annex D The Governance of the MPAI Ecosystem (Informative) 1](#_Toc126691723)

# Introduction

This document is a collection of Use Cases. enabled by Extended Reality (XR) – the combination of Augmented Reality (AR), Virtual Reality (VR) and Mixed Reality (MR) technologies – and enhanced by Artificial Intelligence (AI) technologies. The word “Venue” is used as a synonym for real and virtual environments.

The currently considered use cases are:

1. eSports Tournament (XRV-EST).
2. Live theatrical stage performance (XRV-TSP).
3. Experiential retail/shopping (XRV-ERS).
4. Collaborative immersive laboratory
5. Immersive art experience.
6. DJ/VJ performance at a dance party.
7. Live concert performance.
8. Experiential marketing/branding.
9. Meetings/presentations.

This document also includes the Functional Requirements of the first 3 Use Cases. They should be considered jointly with the Call for Technologies [] and the Commercial Requirements included in the Framework Licence [].

The Call for Technologies seeks to obtain technologies released in conformity with the Framework Licence that support some or all the Functional Requirements to enable MPAI to develop the XR Venues (MPAI-XRV) Technical Specifications. Annex A provides general information about MPAI. More information about the MPAI process to develop standards can be found at

# A Real-Virtual Interaction Model

An important feature of MPAI-XRV is the strong interaction with – and sometimes even interchangeability of – a Real/Virtual Environment and a Virtual/Real World Environment. The MPAI-XRV model, depicted in *Figure 1,* is helpful to guide the analysis of the MPAI-XRV use cases.

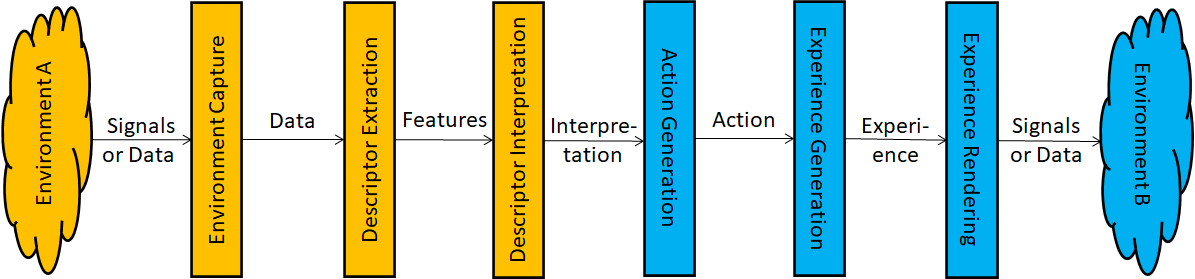


Figure 1 – Environment A (yellow) to Environment B (blue) Interactions

Table 1 defines the functions of the processing elements that capture and process data from a real world or virtual world Environment and deliver an Environment to the same or another virtual world Instance or real world.

Table 1 - The functions of the components in the MPAI-XRV Model

|  |  |
| --- | --- |
| **Environment Capture** | Captures Environment as collections of signals and/or Data. |
| **Feature Extraction** | Analyses Data to extract Descriptors. |
| **Feature Interpretation** | Analyses Descriptors to yield Interpretations. |
| **Action Generation** | Analyses Interpretations to generate Actions. |
| **Experience Generation** | Analyses Actions to generate Environment. |
| **Environment Delivery** | Delivers Environment as collections of signals and/or Data. |

The model assumes that there is a complete symmetry between the actions performed and the data formats exchanged between a real world Environment and a virtual world Environment. Actions initiated by Environment A of Figure 2 has an impact on Environment B and may trigger actions that will influence Environment C.

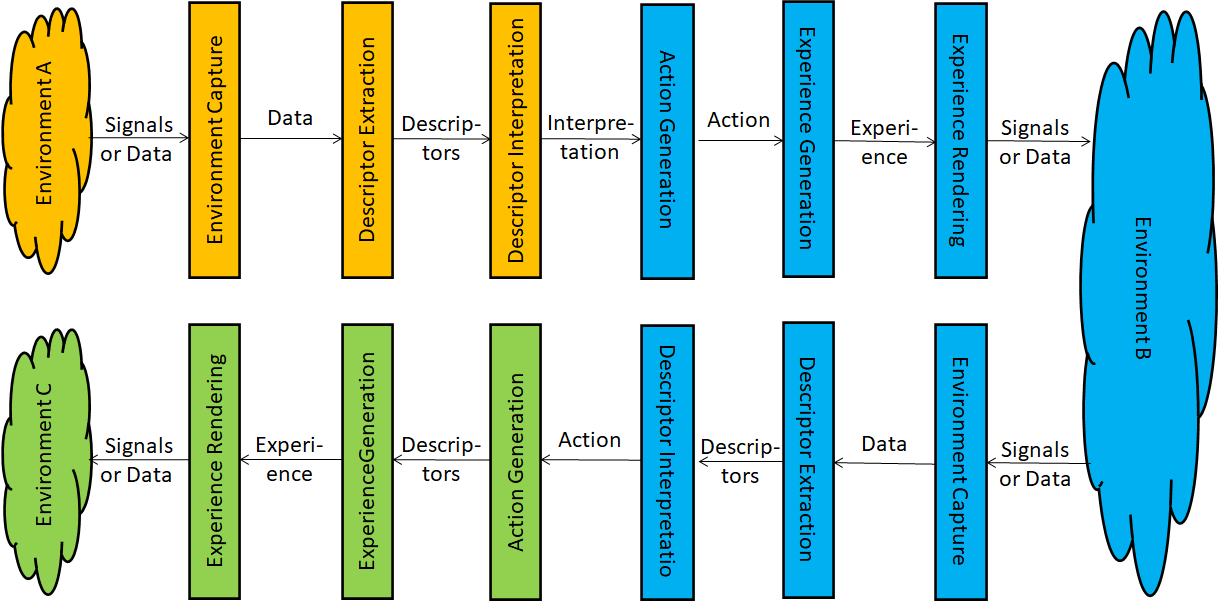


Figure 2 – Environment A to Environment B to Environment C Interactions

In specific Use Cases, it is necessary to introduce new functions called Script and the Participant Data Management function.

|  |  |
| --- | --- |
| Script | A collection of descriptors that the director/producer selects for execution at runtime as a sequence of descriptors controlling the action/experience in both Real and Virtual Worlds. |
| Participants Data Management | The set of legal and ethical rules guiding the acquisition, retention, and processing of Data related to and provided by Participants. |

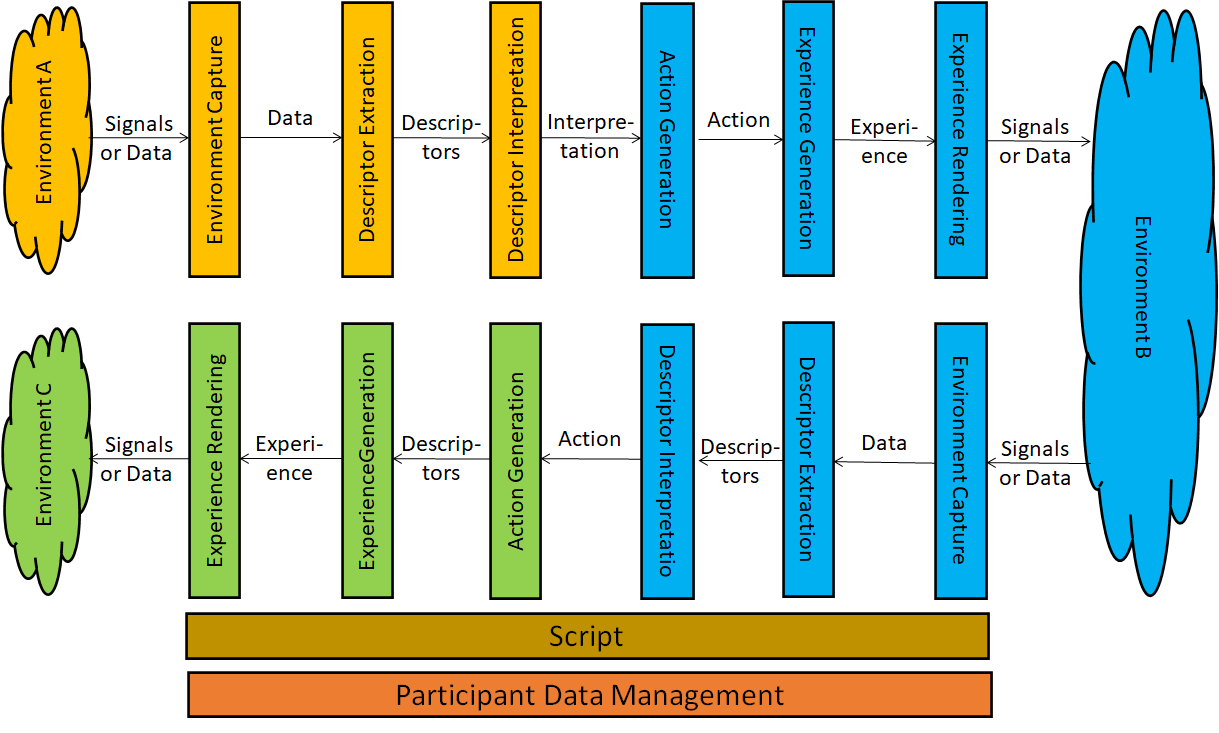


Figure 3 – General MPAI-XRV Reference Model

# Terms and definitions

The meaning of terms used in capital letters in this document have the meaning given in *Table 2*. The Terms of MPAI-wide applicability are defined in Table 7.

*Table 2* *–Terms used in this document*

|  |  |
| --- | --- |
| **Term** | **Definition** |
| Action Generation | The process that triggers a response to the information the machine has acquired and processed, either in the original or in a different Environment, modulating the target Experience. |
| AI Framework | (AIF) The environment where AIWs are executed. |
| AI Module | (AIM) A 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 inputs according to its Function. |
| Avatar | An animated 3D object representing a real or fictitious person in a Virtual Space. |
| Cognitive State | An element of the internal status of a human or avatar reflecting the way they understand the Environment, such as:  For a person: “Confused”, “Confident” and “Assured”.  For a crowd: “my team is going to lose” |
| Data | Information in digital form. |
| * Format | The syntax and semantics of Data. |
| Descriptor | Digital representation of a Feature. |
| * Extraction | The process that extracts Descriptors from Data. |
| * Interpretation | The process that assigns a semantic value to a Descriptor. |
| Emotion | An element of the internal status of a human or avatar resulting from their interaction with the Environment, such as:  For a person, “Victorious”, “Fearful” and “Angry”.  For a group: “Victorious”, “Fearful” and “Disappointed” |
| Environment | A portion of the real or a virtual world sensed by Sensors and affected by Actuators. |
| Extended Reality | (XR) Any of Augmented Reality (AR), Virtual Reality (VR) and Mixed Reality (MR). |
| Feature | An attribute of the Real or Virtual world. |
| Metaverse | A collection of immersive three-dimensional virtual worlds in which people interact as avatars with each other and with software agents, using the metaphor of the real world but without its physical limitations. |
| Participant | A human in an XR Venue. |
| * Data | Data provided by or collected from Participants. |
| * Status | The ensemble of information, expressed by Emotion, Cognitive State and Attitude, derived from observing the collective behaviour of real worldand on-line spectators (via audio, video, interactive controllers, and smartphone apps) in response to actions of a team, a player, or the game. |
| Player Status | The ensemble of information internal to a Player, expressed by Emotion, Cognitive State, and Social Attitude. |
| Social Attitude | An element of the internal status related to the way a human or avatar intends to position vis-à-vis the Environment, e.g.:  For person: “Confrontational”, “Collaborative” and “Aggressive”.  For crowd: “Confrontational”, “Collaborative” and “Aggressive” |
| Use Case | A particular instance of the Application domain target of an MPAI Application Standard. |
| XR Venue | A combination of real or virtual world Environments addressed by MPAI-XRV Use Cases. |

# Use Cases

## Important Information

All MPAI-XRV use cases and many of those not considered here involve the collection of large amounts of potentially sensitive Participant Data. This document does not address the processes that oversee the collection and processing of Participant Data. Thus, document assumes that whatever processing is carried out, it conforms with the necessary and and ethical legal constraints, e.g., with the consent of the right holders of the data[[1]](#footnote-1),[[2]](#footnote-2).

Implementers must take great care in data security, assuring the correct possibility to opt-in/opt-out offered and proper use of the data.

## eSports Tournament (XRV-EST).

### Purpose

To define interfaces between components enabling an XR Theatre (RW) to host any pre-existing VW game for the purpose of producing an esports tournament with RW and VW audience interactivity. To the extent that the game possesses the required interfaces, the XR Theatre can drive action within the VW.

### Description

The eSports Tournament Use Case consists of the following:

1. Two teams of 5 RW players are arranged on either side of a RW stage, each using a computer to compete within a common real-time Massively Multiplayer Online (MMO) VW game space.
2. The 10 players in the VW are represented by avatars each driven by
   1. Role (e.g., magicians, warriors, soldier, etc.).
   2. Properties (e.g., costumes, physical form, physical features).
   3. Actions (e.g., casting spells, shooting, flying, jumping) operating in the VW
3. The VW is populated by
   1. Avatars representing the other players.
   2. Autonomous characters (e.g., dragon, monsters, various creatures)
   3. Environmental structures (e.g., terrain, mountains, bodies of water).
4. The action in the VW is captured by multiple VW cameras and
   1. Projected onto an immersive screen surrounding RW spectators
   2. Live streamed to remote spectators as a 2D video.

with all related sounds of the VW game space.

1. A shoutcaster calls the action as the game proceeds.
2. The image of RW players, player stats or other information or imagery may also be displayed on the immersive screen and the live stream.
3. The RW tournament venue is augmented with lighting and special effects, music, and costumed performers.
4. Interactions:
   1. Live stream viewers interact with one another and with commentators through live chats, Q&A sessions, etc.
   2. RW spectators interact through shouting, waving and interactive devices (e.g., LED wands, smartphones) through processing where:
      1. Data are captured by camera/microphone or wireless data interface (see RW data in *Figure 1*).
      2. Features are extracted and interpreted.
5. RW/VW actions can be generated as a result of:
   1. In-person or remote audience behaviour (RW).
   2. Data collected from VW action (e.g., spell casting, characters dying, bombs exploding)
6. At the end of the tournament, an award ceremony featuring the winning players on the RW stage is held with great fanfare.

## Live theatrical stage performance.

### Purpose

To define interfaces and components to facilitate live multisensory immersive stage performances which ordinarily require extensive on-site show control staff to operate. Use of the XRV allows more direct, precise yet spontaneous show implementation and control to achieve the show director’s vision by freeing staff from repetitive and technical tasks and focusing their artistic and creative skills.

Theatrical stage performances such as Broadway theatre, musicals, dramas, operas, and other performing arts increasingly use video scrims, backdrops, and projection mapping to create digital sets rather than constructing physical stage sets, allowing animated backdrops, and reducing the cost of mounting shows.

The use of immersion domes – especially LED volumes – promises to surround audiences with virtual environments that the live performers can inhabit and interact with the virtual environment and the participants.

Use of AI in these XRV systems will allow:

1. Orchestration of the complex lighting, video, audio, and stage set cues that must adapt to the pace of live performers without extensive staff.
2. Large shows to tour to smaller venues that otherwise could not support complex productions.
3. Virtual-world and real-world attendance and audience interactivity.
4. A more direct connection between the artist and participants by consolidating many complex experiential modalities into a simple user interface.
5. Artists to access a large amount of data from opted-in individuals and incorporate it into the visual and musical performance. Each show can thus be unique for each audience.

### Description and flow of actions

The typical set up can be described as follows:

1. A physical stage.
2. Lighting, stage projections, and SFX.
3. Audience (standing or seated) in the dome.
4. Performers on stage, platforms around domes or moving through the audience (immersive theatres).
5. Multisensory experience delivery system (immersive video and spatialised audio, touch, smell).
6. Capture of biometric data from audience and/or performers from wearables, sensors embedded in the seat, remote sensing.
7. A script or cue list describing the show events.
8. Show operator(s).
9. Metaverse that mirrors selected elements of the RW space. For example, performers on the stage are mirrored by digital twins in the metaverse, using:
   1. Capture body motion (MoCap) to animate an avatar.
   2. Keyed 2D image mapped on a plane.
   3. Volumetrically captured 3D images producing photorealistic digital embodiments.
10. The physical stage and set pieces blend seamlessly into the virtual 3D backdrop projected onto the dome that the spectators perceive as a single immersive environment.
11. The floor of the stage itself and set pieces on the stage may be projection-mapped or wrapped with LED to integrate them into the immersive environment. This allows, for instance set pieces such as a tree, to come alive with moving leaves, blooming flowers or ripening fruits, and for the tree to cast a virtual shadow across the stage from a virtual light source appearing on the dome. Many of these elements may be extracted from the metaverse and projected into the real-world immersive environment.
12. Real performers enter the stage. As they move about the stage, whether dancing, acting, etc., their performance may be mirrored in the metaverse. The performance is accompanied by music, lighting, and SFX.
13. In addition, virtual performers in the metaverse may be projected onto the real-world immersive environment.
14. Script: guides and synchronises the actions of all AIMs as the show evolves from scene to scene.
15. Behaviour: performer’s motion, gesture, vocalisation, biometrics

## Experiential retail/shopping.

### Purpose

To define components and interfaces to facilitate a retail shopping experience enhanced using immersive/interactive technologies driven with AI.

Enhancements includes:

1. Faster locating of products
2. Easy access to product information and reviews.
3. Delivery if special offers
4. Collaborative shopping (members of a group know what other members have purchased)
5. Product annotation according to user preference and theming of the environment according to season and user preferences.
6. Analytics of data collected to inform sales and marketing decisions, inventory control and business model optimisation.
7. Offering remote shoppers the ability to enter a digital twin of real world store as an avatar (as a 3D Graphics or as a volumetric “hologram”) and interact with friends who are physically or virtually present in the real world store.

### Description and flow of actions

The environment displays the following features:

1. It gives the user the impression that it is intelligent because the environment has access to the user’s identity/behaviour/preferences/shopping history/shopping list and is capable to guide the buyer to the area containing products of their supposed interest, propose products, annotate products and to display a particular product and make it flash because the environment thinks it is of interest to the buyer.
2. It broadcasts music etc. to all buyers in the environment driven by the preferences. Friends in the shop at the same time can “meet”, but buyers can opt out from being discoverable (by the store, by friends etc.). Buyers can opt out from the loyalty card and not have the product they buy recorded by the shop.
3. It can be digitally rethemed for different occasions.
4. It offers experience that can takes shape can be anywhere, e.g., in a vehicle or in a public transit space.
5. It enables remote shoppers to virtually enter a digital twin of the store and interact with friends who are physically present in the store for a collaborative shopping experience.

## Collaborative immersive laboratory

### Purpose

Create a collaborative immersive environment allowing citizen scientists and researchers to join physically or virtually via avatar or volumetric representation of themselves for navigation, inspection, analysis, and simulations of scientific or industrial 3D/spatial models/datasets ranging from microscopic to macroscopic.

Examples are:

* View data in its actual 3D or 4D (over time) form through Immersive Reality.
* Present very large data sets that are generated by microscopes, patient, and industrial scanners.
* Format/reformat, qualify, and quantify sliced dataset with enhanced visualisation and analysis tools or import results for rapid correction of metadata for volumetric import.
* Provide tools for investigators to understand complex data sets completely and communicate their findings efficiently.

Objective of an exemplary case: to define interfaces of AI Modules that create 3D models of the fascia from 2D slices sampling microscopic medical images, classify cells based on their spatial phenotype morphology, enable the user to explore, interact with, zoom in the 3D model, count cells, and jump from a portion of the endoderm to another.

### Description

There is a file containing the digital capture of 2D slices, e.g., of the endocrine system.

An AIM reads the file and creates the 3D model of the fascia.

Another AIM finds the cells in the model and classifies them.

A human

1. navigates the 3D model.
2. interacts with the 3D model.
3. zooms in the 3D model (e.g., x2000).
4. converts a confocal image stack into a volumetric model.
5. Analyses the movement of an athlete for setting peak performance goals.

Relevant data formats are:

1. Image Data: TIFF, PNG, JPEG, DICOM, VSI, OIR, IMS, CZI, ND2, and LIF files
2. Mesh Data: OBJ, FBX, and STEP files
3. Volumetric Data: OBJ, PLY, XYZ, PCG, RCS, RCP and E57[[3]](#footnote-3)
4. Supplemental Slides from Powerpoint/Keynote/Zoom
5. 3D Scatterplots from CSV files

### Microscopic dataset visualisation

1. Deals with different object types, e.g.:
   1. 3D Visual Output of a microscope.
   2. 3D model of the brain of a mouse.
   3. Molecules captured as 3D objects by an electronic microscope.
2. Create and add metadata to a 3D audio-visual object:
   1. Define a portion of the object – manual or automatic.
   2. Assign physical properties to (different parts) of the 3D AV object.
   3. Annotate a portion of the 3D AV object.
   4. Create links between different parts of the 3D AV object.
3. Enter, navigate and act on 3D audio-visual objects:
   1. Define a portion of the object – manual or automatic.
   2. Count objects per assigned volume size.
   3. Detect structures in a (portion of) the 3D AV object.
   4. Deform/sculpt the 3D AV object.
   5. Combine 3D AV objects.
   6. Call an anomaly detector on a portion with an anomaly criterion.
   7. Follow a link to another portion of the object.
   8. 3D print (portions of) the 3D AV object.

### Macroscopic dataset visualisation and simulation

1. Deals with different dataset types, e.g.:
   1. Stars, 3D star maps (HIPPARCOS, Tycho Catalogues, etc.).
   2. Deep-sky objects (galaxies, star clusters, nebulae, etc.).
   3. Deep-sky surveys (galaxy clusters, large-scale structures, distant galaxies, etc.).
   4. Satellites and man-made objects in the atmosphere and above, space junks, planetary and Moon positions.
   5. Real-time air traffic.
   6. Geospatial information including CO2 emission maps, ocean temperature, weather, etc.
2. Simulation data
   1. Future/past positions of celestial objects.
   2. Stellar and galactic evolution.
   3. Weather simulations.
   4. Galaxy collisions.
   5. Black hole simulation.
3. Create and add metadata to datasets and simulations:
   1. Assign properties to (different parts) of the datasets and simulations.
   2. Define a portion of the dataset – manual or automatic.
   3. Annotate a portion of the datasets and simulations.
   4. Create links between different parts of the datasets and simulations.
4. Enter, navigate, and act on 3D audio-visual objects:
   1. Search data for extra-solar planets.
   2. Count objects per assigned volume size.
   3. Detect structures and trends in a (portion of) the datasets and simulations.
   4. Call an anomaly detector on a portion with an anomaly criterion.

### Educational lab

1. Experiential learning models simulations for humans.
2. Group navigation across datasets and simulations.
3. Group interactive curricula.
4. Evaluation maps.

### Collaborative CAD

1. Building information management.
2. Collaborative design and art.
3. Collaborative design reviews.
4. Event simulation (emergency planning etc.).
5. Material behaviour simulation (thermal, stress, collision, etc.).

## Immersive art experience.

### Purpose

Define interfaces and components to enhance magical Environments created by skilled artists to provide each user with a unique interactive experience including the ability to modify the environment per their personal style and preferences.

### Description

Immersive art experiences such as Immersive Van Gogh provide visitors with a visually and aurally immersive experience, often based on the work of a specific artist. These are typically passive walk-through and sit-down experiences. The addition of AI to these Environments allows numerous enhancements including the recognition of individual visitors, allowing them to interact with and modify these environments based on pre-selected preferences and style choices. AI style transfer allows the featured artist’s style to be applied to unique visitor interactions which might include AI voice or text-based image diffusion, gesture-based interactions, proximity effects and more. The addition of AR glasses allows visitors to experience, create and interact with “holograms” within the Environment. Biometric wearables allow the AI to monitor and adjust the multisensory experience to maximize target brain/nervous system states related to well-being, restorative states and more. The XR Venue model also allows visitors in the RW and VW to interact.

## DJ/VJ performance at a dance party.

### Purpose

Define interfaces and components to enhance the overall experience within a nightclub, lounge or dance party Environment. The goal is to empower the DJ/VJ to create and control entertaining immersive and interactive experiences that reduce social inhibitions, encourage play, invoke a greater sense of togetherness, encourage personal connections, evoke altered states of consciousness, amplify user’s self-expression and generally create a highly pro-social experience for participants.

### Description

Dance parties, lounges, clubs, and electronic music festivals use powerful visuals, sound and other effects to captivate participants. The DJ (disc jockey) mixes audio tracks, energizes the crowd and is central to the experience. However, the visual artist or VJ (video jockey) is also an important contributor, often supported by lighting, laser and effects operators, dancers, performers and more. Quite often these venues offer peripheral activities as well to further engage participants off of the dance floor, including interactive screens, spatial art, vendors offering costumes and LED accessories. These venues can be thought of as play spaces. Pro-social intoxicants such as alcohol are sometimes used to lower inhibitions that would otherwise limit social connections. XR Venues can supercharge the dance party experience by providing powerful immersive visuals and by including VW participants Assisted by AI, all music, visuals, lights, and effects can be controlled by a single DJ (or immersive jockey) using gestures, simple control surfaces, vocal commandsm and such. In addition, expanded peripheral activities for deeper engagement might include immersive visuals that respond to emergent crowd behaviours, “photonic go-go booths” that modulate immersive visuals to amplify the creative expression of dancers’ movements, and AI-based matchmaking that fosters connections between like-minded attendees.

## Live concert performance.

### Purpose

Define interfaces and components to enhance live musical concerts with AI-driven visuals and special effects and allow enhanced audience participation while extending concert performances into the metaverse.

### Description

Similar to live theatrical stage performances, musical concerts – whether orchestral or popular music – are increasingly using visuals and other effects to enhance the audience experience. A band or orchestral musicians on stage can be substantially enhanced by video projections from a live VJ, audio responsive visuals, image magnification from cameras and other effects. In addition, skilful live mixing of audio is critical to the audience experience, but it complicated by architectural properties of the physical venue. AI can dynamically optimize the listening experience and allow tight synchronization of visuals with spontaneous musical performances in addition to optimizing the VW experience for remote attendees.

## Experiential marketing/branding.

### Purpose

Define interfaces and components to enhance a wide range of experiences in support of corporate branding.

### Description

Wherever there are a lot of people gathered we often find advertisers or corporate brands seeking visibility. Experiential marketing goes beyond simple advertising or signage by offering memorable experiences to attendees. Experiential marketing often makes use of pop-up venues or storefronts co-located at festivals, sporting events, concerts and more. Digital interactive or immersive experiences are increasingly employed, often incorporating branded story-worlds or iconic brand elements. The XRV allows delivery of a unique experience to each participant and deeper engagement to build brand loyalty. In addition, the experience can be extended into the VW to reach a larger number of attendees.

## Meetings/presentations.

### Purpose

Define interfaces and components to enhance live presentations and dialog, both in RW and VW, using rich multimedia, dialog mapping, AI-based mediation and fact checking.

### Description

Meetings and presentations are increasingly hybrid, including both live and virtual attendees, allowing the sharing of rich multimedia content including documents, videos and website links. Use of an XRV for presentations and especially dialog - including political discourse - presents an opportunity for AI to monitor, track, organize and summarize numerous data in real-time to overlook hyperbole and guide the conversation toward rapid convergence on positive outcomes. Real-time fact-finding/fact-checking, dialog mapping (creating a logical tree showing relationships and dependencies between various points raised), group polling and other advanced methods can be employed in an XRV to guide dialog or facilitate presentations.

# Functional requirements

## eSports Tournament (XRV-EST).

### Processing

This subsection applies the diagram represented in Figure 3 to analyse the dta processing performed by XRV-EST.

#### Data/Features capture

Data may come from:

1. The VW
   1. Player/Autonomous character behaviour
      1. Position/orientation
      2. Jumping, walking, flying
      3. Fighting, driving
      4. Casting spells
   2. The environment (not static, evolves with the game)
      1. Terrain
      2. Foliage
      3. Object (building, statue, …)
   3. Players’ viewports.
   4. Metadata: Current game score, players’ status (gas left in car, # of spells left)
2. The RW
   1. Real players
      1. Controller data
      2. Video feeds (face)
      3. Utterances, vocalisation
   2. Spectator (local and remote) response fans supporting their favourite players’ teams
      1. Audio: fans screaming, clapping hands
      2. Video: waving wands, waving arms
      3. Smartphone: interaction with application
      4. Remote: IM chats, web interactives
   3. Other data:
      1. Players’ and teams’ history
      2. Video clips of players

#### Feature extraction

1. Players
   1. Facial descriptors
   2. Audio descriptors: volume, pitch
2. Spectators
   1. Video: descriptors for people moving up and down, waving
   2. Audio descriptors: volume, cheering, booing

#### Feature Interpretation

1. Real Players
   1. Interpret mood, attitude, emotion, sentiment, intention
2. Spectators
   1. Sentiment, intent, and mood of the chats (dynamic)
   2. Warnings to players
3. Other data
   1. Prediction, interpretation of importance based on history

#### Actions

1. Virtual World
   1. Spectators drop an obstacle in front of the opposing team.
   2. A spectator camera (virtual camera controlled by AI) adapts to the environment where it shoots a scene
2. Real World
   1. Visual/auditory cues for the visual/auditory impaired
   2. Modulation of sound, light, and special effects (fireworks, fog, …)
   3. Selection of video feeds to display (local and remote)

### Operations

1. Extraction and interpretation of players’ features, represented by the Gaming Personal Statuses of the players as perceived by a machine.
2. Extraction and interpretation of emergent crowd features, represented by the Crowd Status as perceived by a machine.
3. Extraction and interpretation of the importance of game action based on player/team history, game score, player status.
4. Based on 2-3-4, generation of such actions in the RW as modulation of sound, light, and special effects (fireworks, fog, …) and selection of video feeds to display (local and remote)
5. Based on 3-4 generation of such actions in the virtual world as modulation of the virtual environment including bad weather based on crowd status, control the behaviour of autonomous characters, drop obstacles on the opposing team, in general the fans’ emotional or physical responses conspire to help their team or obstruct the opposing team.
6. Based on 4 control the position and orientation of multiple VW cameras that are displayed both in the RW and live stream.
7. Based on #4 control the display of game stats, player stats and other relevant info in the RW and live stream.
8. Control of autonomous characters, environmental features (mountains, time of day etc.) and other VW actions based on crowd status (VW and live streamed). – not an AIM.

### Reference architecture

Figure 4 provides the Reference Model of the eSports Tournament Use Case.

Diagram

Description automatically generated

Figure - Reference Model of eSports Tournament

### AI Modules

Table 3 gives the list of AIMs, their functions, and their input/output data.

Table 3 - AIMs, their functions and I/O data

|  |  |  |  |
| --- | --- | --- | --- |
| **AIM name** | **Function** | **Input** | **Output** |
| Player Status | Extract Player Status | * Audio * Video * Controllers | * Player Status |
| Participants Status | Extract Participants Status | * Audio * Video * Controllers * Apps | * Participants Status |
| Game State | Creates Game State | * Player location in VW * Player action in VW * Game score * Clock (e.g., timeout) | * Game State |
| Game Action Status | Assesses importance of game action in play | * Nx Player Statuses * Nx Player histories * Team history * Tournament level * Game State | * Game Action Status |
| Audio/Video/SFX generation (RW) | Creates RW experience for spectators (RW/live streaming) | * Player Status * Participants Status * Game Action Status | * Modulated sound * Modulated light * SFXs * Video for display |
| Audio/Video/SFX generation (VW) | Creates VW experience | * Player Status * Participants Status * Game Action Status | * E.g., bad weather if crowd is angry * Drop obstacles on opposing team * Autonomous characters behaviour control |
| VW camera Orientation | Issues commands orientating VW cameras | * Game Action Status | * Orientation commands of multiple VW cameras |

### AIM I/O Data Formats

Table 4 records comments on the data formats.

Table 4 - Comments on data formats

|  |  |
| --- | --- |
| **Data format** | **Comments** |
| Controller data | The data generated by the controllers of the players (console or game controller) or the spectators (wands, or other input devices). Game specific. |
| Player Location (VW) | The location of the player’s avatar in the VW. In other use cases we have introduced Player Attitude (location and orientation). This can be standard. |
| Players’ Action | One of the possible actions of the players’ avatar in the VW. Game specific. |
| Game State | The set of information elements of a game at a given point in time, such as, player position and orientation, player resources, player actions, game levels and score. Game specific |
| Participants Status | The ensemble of information, expressed by Emotion, Cognitive State and Attitude, derived from observing the collective behaviour of RW and on-line spectators (via audio, video, interactive controllers, and smartphone apps) in response to actions of a team, a player, or the game. |
| Player Status | The ensemble of information internal to the Player, expressed by Emotion, Cognitive State, and Attitude. |
| Game Action Status | Importance, priority of Game State in context of team/player history, tournament status. Game specific |
| Action on Environment | Set of developer-defined parameters identifying the action. Game specific. |
| Action on Autonomous Char. | Game specific. |
| Camera orientation commands | Affecting the 6D camera position/orientation, zoom, focus, aperture |

## Live theatrical stage performance

### Operation

A comprehensive view of the devices and data relevant to the XR Theatre Use Case are provided by Figure 5*.*

Diagram

Description automatically generated

Figure 5 - Devices and data of the XR Theatre Use Case

VJ/DJ/Console operator (real-time show control team) inputs command into systems via:

1. Buttons
2. Joysticks
3. Controllers (inertial, proximity, consoles etc.)
4. Hand gestures
5. Headset

The commands into the Action Generator Engine (AGE) modify the multisensory elements (visual, audio, …) of both the Real and the Virtual World Experiences, including:

1. Video/audio clips from media servers.
2. Real-time 2D/3D generated graphics/audio (from a game engine, AI engines generating 3D).
3. Additional video/audio sources
4. 3D objects, characters, scenes.
5. Lighting and special effects.
6. Other experiential elements in both the Real and Virtual Worlds.

Multisensory elements may further be modified by additional real-world and virtual-world data, including:

1. Audience behaviour.
2. Performer behaviour.
3. Events on the stage/dome/metaverse.

### Reference Architecture

Figure 6 provides the Reference Model of the Live theatrical stage performance Use Case.

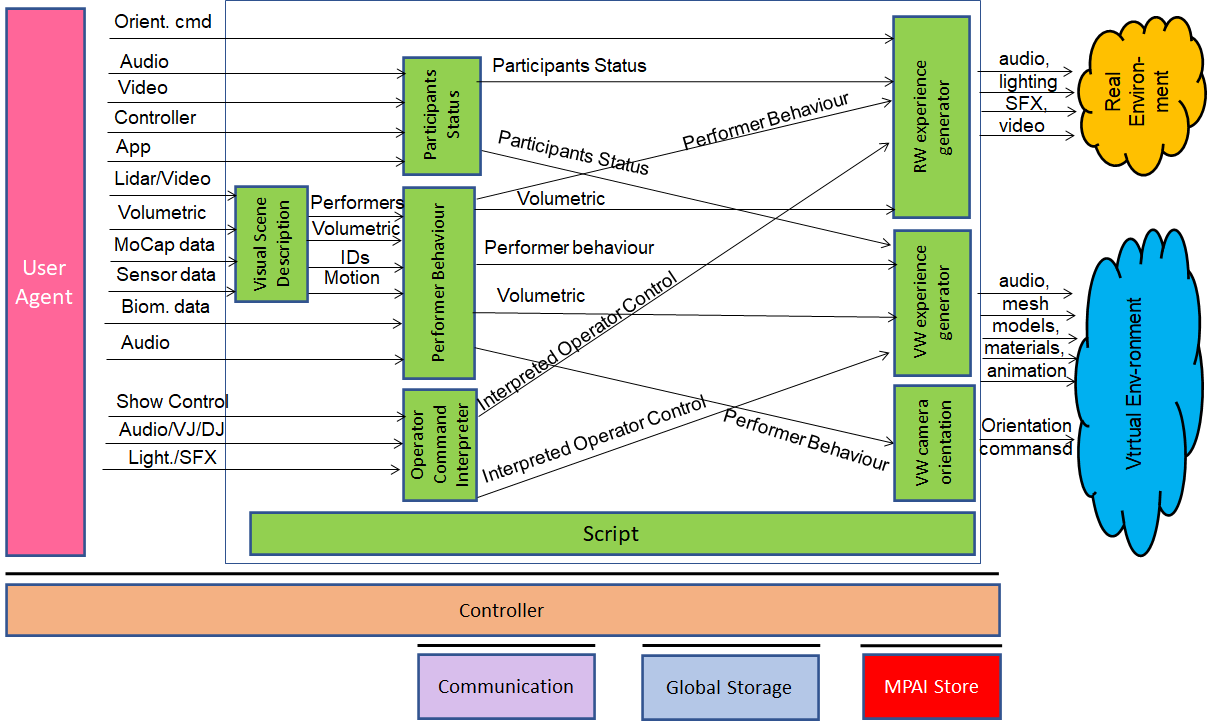


Figure 6 - Reference architecture of Live theatrical stage performance

### AI Modules

Table 3 gives the list of AIMs, their functions and their input/output data

Table 5 - AIMs, their functions and I/O data

|  |  |  |  |
| --- | --- | --- | --- |
| **AIM name** | **Function** | **Input** | **Output** |
| Visual Scene Description | (Feature extraction) Separate and track the individual performers | * Lidar/Video of the stage * MoCap data * Volumetric data * Sensor data | * Performer object * IDs * Movements |
| Performer’s behaviour | Interpret performer movement to create a digital twin and control environment | (Performer’s)   * Performer object & IDs * Biometric data * Audio * Movements | * Performer behaviour. * Performer’s digital twin * Volumetric data |
| Operator command interpreter | Interprets operator commands | * Operator consoles (DJ consoles VJ, lighting and show control console) | * Interpreted operator commands |
| Participants Status | Extract Participants Status | * Audio * Video * Controllers * Apps | * Participants Status |
| RW experience generation | Creates RW experience for spectators (RW/live streaming) | * Performer behaviour * RW Participant Status * Script * Interpreted operator commands * Volumetric data | * Modulated sound * Modulated light * SFXs * Video for display (dome, stage, set pieces, scrims, hologram effects, AR glasses) |
| VW experience generation | Creates VW experience | * Performer behaviour * VW Participant Status * Script * Interpreted operator commands * Volumetric data | * Modulated media in 3D environment (audio, mesh models, materials, animation) |
| VW camera Orientation | Issues commands orientating VW cameras | * Performer behaviour * Script | * Orientation commands of multiple VW cameras for live streaming |

### AIM I/O Data Formats

Table 6records comments on the data formats.

Table 6 - Comments on data formats

|  |  |
| --- | --- |
| **Data formats** | **Comments** |
| MoCap | https://en.wikipedia.org/wiki/List\_of\_motion\_and\_gesture\_file\_formats |
| Volumetric data | FBX  https://www.hhi.fraunhofer.de/en/departments/vca/research-groups/multimedia-communications/research-topics/volumetric-video-formats.html |
| Lidar | Rokoko (https://www.rokoko.com/) |
| Sensor data |  |
| * Accelerometer | D2x,D2y,D2z |
| * Positional tracker | x,y,z |
| * Object tracker | x,y,z |
| Performer behaviour | Gestures (waving, pointing, etc.), Choreography, Mudra, Dance notation  https://en.wikipedia.org/wiki/Dance\_notation BPNM?  <https://www.semanticscholar.org/paper/A-gestural-media-framework-%3A-tools-for-expressive-Jessop/f9883b33802baef8dcd45d6304224c16567a5cf2>  An Automated Structural Approach to Support Theatrical Performances by Introducing Gesture Recognition to a Cuing System  https://www.researchgate.net/publication/335436965\_An\_Automated\_Structural\_Approach\_to\_Support\_Theatrical\_Performances\_by\_Introducing\_Gesture\_Recognition\_to\_a\_Cuing\_System |
| Operator consoles | Sliders, knobs, etc.  Gesture recognition |
| Operator consoles commands | DMX, MIDI, RS-232/422, ethernet  Show control commands generated by interpreting Operator consoles commands potentially based on the script |
| Controllers | Wand, control boxes on the seats, vision system capturing audience motion |
| Participants Status | The ensemble of information, expressed by Emotion, Cognitive State and Attitude, derived from observing the collective behaviour of RW and on-line participants (via audio, video, interactive controllers, and smartphone apps) in response to the theatrical experience. |
| Script | Sequential list of directions for sound, lighting, set movement, performer actions, SFX based on time or events |
| SFX | Fog, strobes, pyrotechnics, stage robotics, motion-base |
| Video for display | In the form of pixels for dome, stage, set pieces, scrims, hologram effects, AR glasses. |
| Camera orientation commands | Affecting the 6D camera position/orientation, zoom, focus, aperture |

## Experiential retail/shopping

### Reference Architecture

Figure 7 provides the Reference Model of the Experiential retail/shopping Use Case. Please note that the the Use Case draws data from the real and virtual store (Environment) and injects data into the same.



Figure 7 - Reference architecture of Experiential retail/shopping

### AI Modules

|  |  |  |  |
| --- | --- | --- | --- |
| **AIM name** | **Function** | **Input** | **Output** |
| Scene Description | (Feature extraction) Separate and track the individual Shoppers, creates digital twins | * Lidar/Video around the store * Volumetric data * RFID, loyalty card, … | * Scene Descriptors * Shopper IDs * Shopper’s Biometr. data * Audio (verbal cmd) * Avatars |
| Shopper status | Extracts the real or virtual shopper status | * Scene Description (RW) * DT (VW) | * Shopper’s status |
| Shopper’s needs | Interprets each shopper data to determine their needs | * Shopper history * Shopper’s list * Shopper’s stated preferences * Shopper Status | * Shopper’s Needs. |
| Shopper experience generation (RW) | Enhance the Shopper’s experience by modifying the environment | * Shopper Status * Shopper’s Needs | * Modulated sound * Modulated light * AR experience * Data to app |
| Shopper experience generation (VW) | Creates VW experience | * Shopper’s Status * Shopper’s Needs * Volumetric/MoCap data | * VR experience * Digital twin of store and shoppers |

### AIM I/O Data Formats

|  |  |
| --- | --- |
| Shopper Object | The DT of a Shopper. |
| Scene Description | Describes the Objects in the Shop including their Spatial Attitudes:  It includes:  The DT of the store (i.e., shelves, the products) and their metadata.  The DT of the Shopper and their Metadata. |
| Shopper history | The collection of shopping experiences, i.e., for each coming to the shop:   1. What was purchased. 2. What was considered but not purchased. 3. Shopper’s list. 4. Scene Descriptions. |
| Shopper’s list | The list of immediate preference for the current coming to the shop:   1. Product 2. Priority |
| Shopper’s stated preferences | The list of preferences (predilections) wrt product types, e.g., dietary, health |
| Shopper’s Status | The ensemble of information, expressed by Emotion, Cognitive State and Attitude, derived from observing the Scene Description (via audio, video, smartphone apps, and the VW) in response to actions of a real or virtual Shopper. |
| Shopper’s Needs | The interpretation of the shopper’s movements coupled with list, stated preferences, and history that includes the anticipated needs: an inference based on shopper’s history, list, movements, commands, and stated preferences |
| Shopper’s list | What shopper wants to buy, it is obtained, e.g., from the smartphone app. |

Store operational control oversees and manages all functions of the store, hours of operation, special offers, theming of store based on seasons, time of day etc.

yed in an XRV to guide dialog or facilitate presentations.

# References

## Normative References

MPAI-XRV normatively references the following documents:

1. MPAI; Technical Specification: Artificial Intelligence Framework (MPAI-AIF) V1.1; https://mpai.community/standards/mpai-aif/.
2. MPAI; Technical Specification: Multimodal Conversation (MPAI-MMC) V1.2; https://mpai.community/standards/mpai-mmc/.
3. MPAI; Technical Specification: Context-based Audio Experience (MPAI-MMC) V1.4; https://mpai.community/standards/mpai-mmc/.
4. MPAI; Patent Policy; https://mpai.community/about/the-mpai-patent-policy/

## Informative References

1. MPAI; Technical Specification: The Governance of the MPAI Ecosystem V1, 2021; https://mpai.community/standards/mpai-gme/.
2. MPAI; Technical Report: MPAI Metaverse Model (MPAI-MMM)– Functionalities; V1; https://mmm.mpai.community/.
3. Basics about MPAI

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. Indeed, research has shown that data coding with AI-based technologies is generally *more efficient* than with existing technologies for, e.g., compression and feature-based description.

However, some AI technologies may carry inherent risks, e.g., in terms of bias toward some classes of users. Therefore, the need for standardisation is more important and urgent than ever.

The international, unaffiliated, not-for-profit MPAI – Moving Picture, Audio and Data Coding by Artificial Intelligence Standards Developing Organisation has the mission to develop *AI-enabled data coding standards*. MPAI Application Standards enable the development of AI-based products, applications, and services.

As a part of its mission, MPAI has developed standard operating procedures to enable users of MPAI implementations to make informed decision about their applicability. Central to this is the notion of Performance, defined as a set of attributes characterising a reliable, trustworthy, and ethical implementation.

For the aforementioned reasons, to fully achieve the MPAI mission, Technical Specifications must be complemented by an ecosystem designed, created and managed to underpin the life cycle of MPAI standards through the steps of specification, technical testing, assessment of product safety and security, and distribution.

In the following, Terms beginning with a capital letter are defined in *Table 2* if they are specific to this project and in Table 7 if they are common to all MPAI Standards.

The MPAI Ecosystem is fully specified in [5]. It is composed of:

* MPAI as provider of Technical, Conformance and Performance Specifications.
* Implementers of MPAI standards.
* MPAI-appointed Performance Assessors.
* The MPAI Store which takes care of secure distribution of validated Implementations.
* End Users who use Implementations from the MPAI Store.

*Figure 8* depicts the MPAI-AIF Reference Model under which Implementations of MPAI Application Standards and user-defined MPAI-AIF conforming applications operate.

An AIF Implementation allows execution of AI Workflows (AIW), composed of basic processing elements called AI Modules (AIM).

MPAI Application Standards normatively specify Syntax and Semantics of the input and output data and the Function of the AIW and the AIMs, and the Connections between and among the AIMs of an AIW.

In particular, an AIM is defined by its Function and data, but not by its internal architecture, which may be based on AI or data processing, and implemented in software, hardware or hybrid software and hardware technologies.



Figure 8 - The AI Framework (AIF) Reference Model and its Components

MPAI defines Interoperability as the ability to replace an AIW or an AIM Implementation with a functionally equivalent Implementation. MPAI also defines 3 Interoperability Levels of an AIW executed in an AIF:

*Level 1* – Conforming withthe MPAI-AIF Standard.

*Level 2* – Conforming with an MPAI Application Standard.

*Level 3* – Positively Assessed by a Performance Assessor.

MPAI offers Users access to the promised benefits of AI with a guarantee of increased transparency, trust and reliability as the Interoperability Level of an Implementation moves from 1 to 3. Additional information on Interoperability Levels is provided in Annex D.

1. MPAI-wide terms and definitions

The Terms used in this standard whose first letter is capital and are not already included inTable 1are defined in Table 7*.*

Table 7 - 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 Module (AIM) | A 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 inputs according to its Function. |
| AIF Metadata | The data set describing the capabilities of an AIF set by the AIF Implem­enter. |
| AIM Metadata | The data set describing the capabilities of an AIM set by the AIM Implem­enter. |
| Application Programming Interface (API) | A software interface that allows two applications to talk to each other |
| Application Standard | An MPAI Standard specifying AIWs, AIMs, Topologies and Formats suitable for a particular application domain. |
| Channel | A physical or logical connection between an output Port of an AIM and an input Port of an AIM. The term “connection” is also used as a synonym. |
| Communication | The infrastructure that implements message passing between AIMs. |
| Component | One of the 9 AIF elements: Access, AI Module, AI Workflow, Commun­ication, Controller, Internal Storage, Global Storage, MPAI Store, and User Agent. |
| 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 | Information in digital form. |
| Data Format | The standard digital representation of Data. |
| Data Semantics | The meaning of Data. |
| Device | A hardware and/or software entity running at least one instance of an AIF. |
| 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. |
| Event | An occurrence acted on by an Implementation. |
| 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. |
| 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). |
| Internal Storage | A Component to store data of the individual AIMs. |
| Interoperability | The ability to functionally replace an AIM/AIW with another AIM/AIW having the same Interoperability Level |
| Interoperability Level | The attribute of an AIW and its AIMs to be executable in an AIF Implementation and to be:   1. Implementer-specific and satisfying the MPAI-AIF Standard *(Level 1)*. 2. Specified by an MPAI Application Standard (*Level 2)*. 3. Specified by an MPAI Application Standard and certified by a Performance Assessor (*Level 3)*. |
| Knowledge Base | Structured and/or unstructured information made accessible to AIMs via MPAI-specified interfaces |
| Message | A sequence of Records. |
| 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 |
| Port | A physical or logical communication interface of an AIM. |
| 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 | Data 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. |
| Specification | A collection of normative clauses. |
| 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:   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. |
| Testing Laboratory | A laboratory accredited by MPAI to Assess the Grade of Performance of Implementations. |
| Time Base | The protocol specifying how AIF 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. |

1. Notices and Disclaimers Concerning MPAI Standards (Informative)

The notices and legal disclaimers given below shall be borne in mind when [downloading](https://www.mpai.community/resources/) and using approved MPAI Standards.

In the following, “Standard” means the collection of four MPAI-approved and [published](https://www.mpai.community/resources/) documents: “Technical Specification”, “Reference Software” “Conformance Testing” and, where applicable, “Performance Testing”.

The life cycle of MPAI Standards

MPAI Standards are developed in accordance with the [MPAI Statutes](https://mpai.community/statutes/). An MPAI Standard may only be developed when a Framework Licence has been adopted. MPAI Standards are developed by specially established MPAI Development Committees that operate based on consensus, as specified in Annex 1 of the [MPAI Statutes](https://mpai.community/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](https://mpai.community/resources/) to determine the status of any given published MPAI Standard.

Comments 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](mailto:secretariat@mpai.community). 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 of any kind 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 practising 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, a positive Performance Assessment of an implementation applies exclusively in the context of the [MPAI Governance](https://mpai.community/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](mailto:secretariat@mpai.community).

The Reference Software of an MPAI Standard is released with the [MPAI Modified Berkeley Software Distribution licence](https://mpai.community/about/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.

1. The Governance of the MPAI Ecosystem (Informative)

**Level 1 Interoperability**

Concerning *Figure 8* MPAI issues and maintains a standard – called MPAI-AIF – whose components are:

1. An environment called AI Framework (AIF) running AI Workflows (AIW) composed of inter­connected AI Modules (AIM) exposing standard interfaces.
2. A distribution system of AIW and AIM Implementation called MPAI Store from which an AIF Implementation can download AIWs and AIMs.

A Level 1 Implementation shall implement the MPAI-AIF Technical Specification executing AIWs composed of AIMs able to call the MPAI-AIF APIs.

|  |  |
| --- | --- |
| Implementers’ benefits | Upload to the MPAI Store and have globally distributed Implementations of   * AIFs conforming to MPAI-AIF. * AIWs and AIMs performing prop­rietary functions executable in AIF. |
| Users’ benefits | Rely on Implementations that have been tested for security. |
| MPAI Store’s role | * Tests the Conformance of Implementations to MPAI-AIF. * Verifies Implementations’ security, e.g., absence of malware. * Indicates unambiguously that Implementations are Level 1. |

**Level 2 Interoperability**

In a Level 2 Implem­entation, the AIW must be an Implementation of an MPAI Use Case and the AIMs must con­form with an MPAI Applicati­on Standard.

|  |  |
| --- | --- |
| Implementers’ benefits | Upload to the MPAI Store and have globally distributed Implementations of   * AIFs conforming to MPAI-AIF. * AIWs and AIMs conforming to MPAI Application Standards. |
| Users’ benefits | * Rely on Implementations of AIWs and AIMs whose Functions have been reviewed during standardisation. * Have a degree of Explainability of the AIW operation because the AIM Func­tions and the data Formats are known. |
| Market’s benefits | * Open AIW and AIM markets foster competition leading to better products. * Competition of AIW and AIM Implementations fosters AI innovation. |
| MPAI Store’s role | * Tests Conformance of Implementations with the relevant MPAI Standard. * Verifies Implementations’ security. * Indicates unambiguously that Implementations are Level 2. |

**Level 3 Interoperability**

MPAI does not generally set standards on how and with what data an AIM should be trained. This is an important differentiator that promotes competition leading to better solutions. However, the performance of an AIM is typically higher if the data used for training are in greater quantity and more in tune with the scope. Training data that have a large variety and cover the spec­trum of all cases of interest in breadth and depth typically lead to Implementations of higher “quality”.

For Level 3, MPAI normatively specifies the process, the tools and the data or the characteristics of the data to be used to Assess the Grade of Performance of an AIM or an AIW.

|  |  |
| --- | --- |
| Implementers’ benefits | May claim their Implementations have passed Performance Assessment. |
| Users’ benefits | Get assurance that the Implementation being used performs correctly, e.g., it has been properly trained. |
| Market’s benefits | Implementations’ Performance Grades stimulate the development of more Performing AIM and AIW Implementations. |
| MPAI Store’s role | * Verifies the Implementations’ security * Indicates unambiguously that Implementations are Level 3. |

**The MPAI ecosystem**

*Figure 9* is a high-level description of the MPAI ecosystem operation applicable to fully conforming MPAI implementations as specified in the Governance of the MPAI Ecosystem Specification [5]:

1. MPAI establishes and controls the not-for-profit MPAI Store.
2. MPAI appoints Performance Assessors.
3. MPAI publishes Standards.
4. Implementers submit Implementations to Performance Assessors.
5. If the Implementation Performance is acceptable, Performance Assessors inform Implementers and the MPAI Store.
6. Implementers submit Implementations to the MPAI Store tested for Confor­mance and security.
7. Users download and use Implementations and submit experience scores.

*Diagram

Description automatically generated*

Figure 9 - The MPAI ecosystem operation

1. https://openvoicenetwork.org/documents/ovn\_ethical\_guidlines\_voice\_experiences.pdf [↑](#footnote-ref-1)
2. https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html [↑](#footnote-ref-2)
3. https://info.vercator.com/blog/what-are-the-most-common-3d-point-cloud-file-formats-and-how-to-solve-interoperability-issues [↑](#footnote-ref-3)