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

**Public document**

|  |  |
| --- | --- |
| **N1213** | 2023/05/17 |
| **Source** | Requirements (CAV) |
| **Title** | Technical Report - Connected Autonomous Vehicles (MPAI-CAV) WD0.5 |
| **Target** | MPAI Community |

DISCLAIMER

This Technical Report defines the scope of the future Connected Autonomous Vehicle (CAV) standard. It identifies 4 Subsystems and, for each subsystem:

1. The function, interfaces, I/O data format, and AIM topology of the AI Workflow (AIW) implementing the Subsystems.
2. The function, interfaces, and I/O data format functional functionalities and specifications of the AI Modules (AIM).

MPAI plans the following steps:

1. Publish the CAV Technical Report.
2. Receive comments on the Technical Report.
3. Issue one or more Calls for Technologies
4. Develop a Technical Specification.

This document is a work in progress.

|  |  |
| --- | --- |
|  | Moving Picture, Audio and Data Coding  by Artificial Intelligence  [www.mpai.community](http://www.mpai.community) |

**MPAI Technical Report**

**Connected Autonomous Vehicles**

**MPAI-CAV**

|  |
| --- |
| **WD0.5** |

**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 2 - Notices and Disclaimers.

© Copyright MPAI 2021-23. All rights reserved

**Technical Report**

**Connected Autonomous Vehicles**

**V1 (under development)**

[1 Introduction 5](#_Toc134018628)

[2 Scope 7](#_Toc134018629)

[3 Terms and definitions 9](#_Toc134018630)

[4 References 13](#_Toc134018631)

[4.1 Normative References 13](#_Toc134018632)

[5 Human-CAV Interaction (HCI) 13](#_Toc134018633)

[5.1 Subsystem description 13](#_Toc134018634)

[5.2 Reference architecture 13](#_Toc134018635)

[5.3 I/O Data of the HCI AIW 15](#_Toc134018636)

[5.4 I/O Data of the HCI AIMs 16](#_Toc134018637)

[5.5 Data Formats 17](#_Toc134018638)

[5.5.1 HCI-AMS Commands 17](#_Toc134018639)

[5.5.2 AMS-HCI Response 17](#_Toc134018640)

[5.5.3 Audio 17](#_Toc134018641)

[5.5.4 Audio Scene Descriptors 18](#_Toc134018642)

[5.5.5 Visual Scene Descriptors 18](#_Toc134018643)

[5.5.6 Visual Scene Descriptors 18](#_Toc134018644)

[5.5.7 Avatar Model 19](#_Toc134018645)

[5.5.8 Face Descriptors 19](#_Toc134018646)

[5.5.9 Face ID 19](#_Toc134018647)

[5.5.10 Face Object 20](#_Toc134018648)

[5.5.11 FER Commands 20](#_Toc134018649)

[5.5.12 FER View 20](#_Toc134018650)

[5.5.13 Full Environment Representation 20](#_Toc134018651)

[5.5.14 Human Object 20](#_Toc134018652)

[5.5.15 Machine (Face) 21](#_Toc134018653)

[5.5.16 Machine (Gesture) 21](#_Toc134018654)

[5.5.17 Machine (Speech) 21](#_Toc134018655)

[5.5.18 Meaning 21](#_Toc134018656)

[5.5.19 Object ID 22](#_Toc134018657)

[5.5.20 Orientation 23](#_Toc134018658)

[5.5.21 Personal Status 23](#_Toc134018659)

[5.5.22 Physical Objects 24](#_Toc134018660)

[5.5.23 Position 24](#_Toc134018661)

[5.5.24 Speaker ID 24](#_Toc134018662)

[5.5.25 Speech Descriptors 24](#_Toc134018663)

[5.5.26 Speech Object 24](#_Toc134018664)

[5.5.27 Text 25](#_Toc134018665)

[5.5.28 Video 25](#_Toc134018666)

[5.5.29 Visual Scene Descriptors 25](#_Toc134018667)

[6 Environment Sensing Subsystem (ESS) 26](#_Toc134018668)

[6.1 Subsystem description 26](#_Toc134018669)

[6.2 Reference architecture 26](#_Toc134018670)

[6.3 Input and output data 27](#_Toc134018671)

[6.4 AI Modules 28](#_Toc134018672)

[6.5 I/O Data summary 28](#_Toc134018673)

[6.6 Data Formats 29](#_Toc134018674)

[6.6.1 Radar Data 29](#_Toc134018675)

[6.6.2 LiDAR Data 29](#_Toc134018676)

[6.6.3 Camera Data 30](#_Toc134018677)

[6.6.4 Ultrasound Data 30](#_Toc134018678)

[6.6.5 Offline Map Data 30](#_Toc134018679)

[6.6.6 Audio Data 31](#_Toc134018680)

[6.6.7 Microphone Array Geometry 31](#_Toc134018681)

[6.6.8 GNSS Data 32](#_Toc134018682)

[6.6.9 Spatial Attitude 32](#_Toc134018683)

[6.6.10 Other Environment Data 33](#_Toc134018684)

[6.6.11 Scene Description Format 33](#_Toc134018685)

[6.6.12 Object Format 34](#_Toc134018686)

[6.6.13 Lidar Scene Description 35](#_Toc134018687)

[6.6.14 Radar Scene Description 35](#_Toc134018688)

[6.6.15 Road Topology 35](#_Toc134018689)

[6.6.16 Camera Scene Descriptors 36](#_Toc134018690)

[6.6.17 Ultrasound Scene Description 36](#_Toc134018691)

[6.6.18 Maps Scene Description 36](#_Toc134018692)

[6.6.19 Audio Scene Description 36](#_Toc134018693)

[6.6.20 Alert 37](#_Toc134018694)

[6.6.21 Basic Environment Representation 37](#_Toc134018695)

[7 Autonomous Motion Subsystem (AMS) 38](#_Toc134018696)

[7.1 Use Case description 38](#_Toc134018697)

[7.2 Reference architecture 39](#_Toc134018698)

[7.3 Input and output data 40](#_Toc134018699)

[7.4 AI Modules 40](#_Toc134018700)

[7.4.1 Summary of Subsystem data 41](#_Toc134018701)

[7.4.2 Basic Environment Representation 41](#_Toc134018702)

[7.4.3 CAV Identifier 41](#_Toc134018703)

[7.4.4 Commands 42](#_Toc134018704)

[7.4.5 Feedback 42](#_Toc134018705)

[7.4.6 Road State 42](#_Toc134018706)

[7.4.7 Full Environment Representation 43](#_Toc134018707)

[7.4.8 Goal 43](#_Toc134018708)

[7.4.9 Offline map 44](#_Toc134018709)

[7.4.10 Path 44](#_Toc134018710)

[7.4.11 Pose 44](#_Toc134018711)

[7.4.12 Route 44](#_Toc134018712)

[7.4.13 Traffic rules 44](#_Toc134018713)

[7.4.14 Traffic Signals 44](#_Toc134018714)

[7.4.15 Trajectory 45](#_Toc134018715)

[8 Motion Actuation Subsystem (MAS) 45](#_Toc134018716)

[8.1 Use Case description 45](#_Toc134018717)

[8.2 Reference architecture 45](#_Toc134018718)

[8.3 Input and output data 46](#_Toc134018719)

[8.4 AI Modules 46](#_Toc134018720)

[8.4.1 Summary of Motion Actuation Subsystem data 47](#_Toc134018721)

[8.4.2 Accelerometer data 47](#_Toc134018722)

[8.4.3 Brakes Command 47](#_Toc134018723)

[8.4.4 Brakes Feedback 47](#_Toc134018724)

[8.4.5 Command from AMS 48](#_Toc134018725)

[8.4.6 Feedback to AMS 48](#_Toc134018726)

[8.4.7 Spatial Attitude 48](#_Toc134018727)

[8.4.8 Odometer Data 48](#_Toc134018728)

[8.4.9 Other Environment Data 48](#_Toc134018729)

[8.4.10 Road Wheel Direction Command 48](#_Toc134018730)

[8.4.11 Road Wheel Direction Feedback 49](#_Toc134018731)

[8.4.12 Road Wheel Motor Command 49](#_Toc134018732)

[8.4.13 Road Wheel Motor Feedback 49](#_Toc134018733)

[8.4.14 Speedometer 49](#_Toc134018734)

[9 CAV-to-Everything (V2X) 49](#_Toc134018735)

[9.1 Description 49](#_Toc134018736)

[9.2 Input and output data 50](#_Toc134018737)

[9.2.1 CAVs within range 50](#_Toc134018738)

[9.2.2 CAV-aware equipment 51](#_Toc134018739)

[9.2.3 Other non-CAV vehicles 51](#_Toc134018740)

[9.2.4 Pedestrians 51](#_Toc134018741)

[10 Data privacy 51](#_Toc134018742)

[10.1 Human-CAV Interaction (HCI) 51](#_Toc134018743)

[10.2 Environment Sensing Subsystem (ESS) 52](#_Toc134018744)

[10.3 CAV to Everything (V2X) 52](#_Toc134018745)

[10.4 Autonomous motion subsystem (AMS) 52](#_Toc134018746)

[10.5 Motion Actuation Subsystem (MAS) 52](#_Toc134018747)

[Annex 1 - – General MPAI Terminology 53](#_Toc134018748)

[Annex 2 - - Notices and Disclaimers Concerning MPAI Standards (Informative) 56](#_Toc134018749)

[Annex 3 - – The Governance of the MPAI Ecosystem (Informative) 58](#_Toc134018750)

[10.6 Informative References 59](#_Toc134018751)

# Introduction

In recent years, Artificial Intelligence (AI) and related technologies have been introduced in 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 rule, MPAI standards include four documents: Technical Specification, Reference Software Specifications, Conformance Testing Specifications, and Performance Assessment Specifications.

The last type of Specification includes standard operating procedures to enable users of MPAI Implementations to make informed decision about their applicability based on the notion of Performance, defined as a set of attributes characterising a reliable and trustworthy implementation.

In the following, Terms beginning with a capital letter are defined in

*Table 4* if they are specific to this Standard and in *Table 15* if they are common to all MPAI Standards.

In general, MPAI Application Standards are defined as aggregations – called AI Workflows (AIW) – of processing elements – called AI Modules (AIM) – executed in an AI Framework (AIF). 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 AIF that executes an AIW. The AIW and its AIMs may have 3 Levels:

*Level 1 –* Implementer-specific and satisfying the MPAI-AIF Standard.

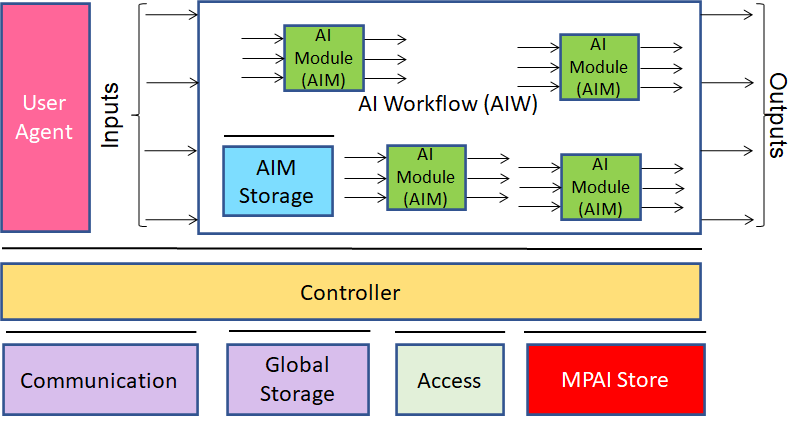
*Level 2 –* Specified by an MPAI Application Standard.

*Level 3 –* Specified by an MPAI Application Standard and certified 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 reference [10].

*Figure 1* depicts the MPAI-AIF Reference Model under which Implementations of MPAI Applic­ation Standards and user-defined MPAI-AIF Conforming applications operate.

MPAI Application Standards normatively specify the 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.



*Figure 1 – The AI Framework (AIF) Reference Model and its Components*

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.

MPAI Standards are designed to enable a User to obtain, via standard protocols, an Implementation of an AIW and of the set of corresponding AIMs and execute it in an AIF Implementation. The MPAI Store in *Figure 1* is the entity from which Implementations are downloaded. MPAI Standards assume that the AIF, AIW, and AIM Implementations may have been developed by independent implementers. A necessary condition for this to be possible, is that any AIF, AIW, and AIM implementations be uniquely identified. MPAI has appointed an ImplementerID Registration Authority (IIDRA) to assign unique ImplementerIDs (IID) to Implementers.[[1]](#footnote-1)

A necessary condition to make possible the operations described in the paragraph above is the existence of an ecosystem composed of Conformance Testers, Performance Assessors, the IIDRA and an instance of the MPAI Store. Reference [10] provides an informative example of such ecosystem.

# Scope

The scope of this Technical Report is to specify:

1. The architecture of a Connected Autonomous Vehicle (CAV), defined a system that:
   1. Ha the capability to autonomously reach a target destination by:
      1. Understanding human utterances.
      2. Planning a Route.
      3. Sensing and building a Representations of the external Environment.
      4. Exchanging such Representations and other Data with other CAVs and CAV-aware entities, such as, Roadside Units and Traffic Lights.
      5. Making decisions about how to execute the Route.
      6. Acting on the CAV motion actuation.
   2. Is structured in four Subsystems as depicted in *Figure 2* and described in *Table 1*:
      1. Human-CAV Interaction (HCI).
      2. Environment Sensing Subsystem (ESS)
      3. Autonomous Motion Subsystem (AMS).
      4. Motion Actuation Subsystem (MAS).



*Figure 2 – The CAV subsystems*

1. Has each Subsystems implemented as an AI Framework (AIF), executing an AI Workflow composed of AI Modules as specified in [1].
2. Supports secure execution according to the MPAI-AIF Technical Specification (MPAI-AIF) V2 [].
3. Provides for each Subsystem:
   1. The Function.
   2. The Functional Requirements of all I/O data.
   3. The Formats of some I/O data.
   4. The Topology of the Components (AI Modules).
4. Provides for each Subsystem *Component* (AI Module):
   1. The Function.
   2. The Functional Requirements of some I/O data.
   3. The Formats of some I/O data.

This Technical Report does not include the mechanical parts of CAV. Reference is made to the interfaces of the Motion Actuation Subsystem with such parts.

*Figure 3* depicts the context and actors of CAV operation and *Figure 2* the internal CAV Subsystem architecture.

A picture containing icon

Description automatically generated

*Figure 3 – An environment of CAV operation*

*Table 1* provides more extended definitions of the Functions performed by the four CAV Subsystems.

*Table 1 – The Functions of the MPAI-CAV Subsystems*

|  |  |
| --- | --- |
| **Subsystem name** | **Function** |
| *Human-CAV Interaction* (HCI) | 1. Recognises the humans having rights to the CAV. 2. Receives and passes to the AMS instructions about the target destination. 3. Interacts with humans by assuming the shape of an avatar. 4. Activate other Subsystems as required by humans. 5. Provides the Full Environment Representation received from the AMS for passengers to use. |
| *Environment Sensing Subsystem* (ESS) | 1. Acquires and processes information from the Environment. 2. Produces the Basic Environment Representation 3. Sends the Basic Environment Representation to the AMS. |
| *Autonomous Motion Subsystem* (AMS) | 1. Computes the Route to destination based on information received from the HCI. 2. Receives the Basic Environment Representation of the ESS and of other CAVs in range. 3. Creates the Full Environment Representation. 4. Issues commands to the MAS to drive the CAV to the intended destination. |
| *Motion Actuation Subsystem* (MAS) | 1. Sends its Spatial Attitude and other Environment information to the ESS. 2. Receives/actuates motion commands in the Environment. 3. Sends feedback to the AMS |

The following high-level workflow illustrates a CAV operation example and the role of CAV Subsystems:

A *human* with appropriate credentials requests the CAV, via *Human-CAV Interaction*, to take the human to a given Pose.

*Human-CAV Interaction* authenticates the human, interprets the request, and passes a command to the *Autonomous Motion Subsystem*. The human may subsequently integrate/correct their instructions.

*Autonomous Motion Subsystem*:

Requests *Environment Sensing Subsystem* to provide the current Pose.

Computes the Route and may offer options to authenticated humans.

*Environment Sensing Subsystem* computes and sends the Basic Environment Representation to the *Autonomous Motion Subsystem.*

*Autonomous Motion Subsystem*:

* 1. Receives the Basic Environment Representations from the Environment Sensing Subsystem
  2. Exchanges the Basic Environment Representation with other CAVs and computes the Full Environment Representation.
  3. Issues appropriate commands to the Motion Actuation Subsystem.

While the CAV moves, the humans in the cabin may:

* 1. Interact and hold conversation with other humans on board and the Human-CAV Interaction Subsystem.
  2. Issue commands.
  3. Requests the Full Environment Representation to view the environment.
  4. Interact with (humans in) other CAVs.

Each of the 4 subsystems is defined for implementation as an instance of the MPAI-defined AI Framework (AIF).

This Technical Report has been developed by the Connected Autonomous Vehicles group of the Requirements Standing Committee. MPAI may publish more versions of this Technical Report and intends to publish a Technical Specification where the AIM and AIW I/O Data Formats will all be specified.

# Terms and definitions

*Table 2* defines the terms used in this document. Terms are organised by the CAV Subsystems identified in *Figure 2* and the CAV for CAV-specific terms. The general MPAI Terms are defined in *Table 15*.

*Table 2 – Definition of Terms used in this document organised by CAV Subsystems*

|  |  |  |
| --- | --- | --- |
| Legend | AMS | Autonomous Motion Subsystem |
|  | CAV | Connected Autonomous Vehicle |
|  | ESS | Environment Sensing Subsystem |
|  | HCI | Human-CAV Interaction |
|  | MAS | Motion Actuation Subsystem |

|  |  |  |
| --- | --- | --- |
| **Sub System** | **Term** | **Definition** |
| AMS | Command | High-level instructions whose execution allows CAV to reach a Goal. |
| AMS | Full Environment Representation (FER) | A description of the Environment using the Basic Environment Representations of the ego CAV and other CAVs. |
| AMS | Goal | The planned Spatial Attitude at the end of a Decision Horizon. |
| AMS | Orientation | The set of the 3 roll, pitch, yaw angles indicating the rotation around the principal axis (x) of an Object, its y axis having an angle of 90˚ counterclockwise (right-to-left) with the x axis and its z axis (pointing up toward the viewer viewing from above). |
| AMS | Path | A sequence of Poses 𝑝𝑖 = (xi, yi, zi, αi, βi, γi) in the Offline Map. |
| AMS | Pose | Position and Orientation of the CAV in the Offline Map p = (𝑥, 𝑦, z, α, β, γ) |
| AMS | Position | The coordinates of the current place of a CAV. |
| AMS | Route | A sequence of Way Points. |
| AMS | Spatial Attitude | CAV’s Position, Orientations and 1st and their 2nd order time derivatives at a given time. |
| AMS | Traffic Rules | The digital representation of the traffic rules applying to a Waypoint. |
| AMS | Trajectory | A sequence of Spatial Attitudes *si* (*s1,s2,…si*) and the expected time each Spatial Attitude will be reached. |
| AMS | Waypoint | A point 𝑤𝑖 given as (𝑥𝑖, 𝑦𝑖, z𝑖) coordinates in an Offline Map. |
| AMS | Platooning | The joint displacement of a group of CAVs along a shared Trajectory. |
| CAV | Accuracy | The degree of confidence of an estimated value (by computation, inference, etc.) |
| CAV | Centre | The point in the CAV selected as its representative, i.e., with coordinates (0,0,0). |
| CAV | Compatibility | The ability of a data format to be converted to another data format without loss of information |
| CAV | Connected Autonomous Vehicle | A vehicle able to autonomously reach a Pose by:   1. Understanding human utterances in the Subsystem (HCI). 2. Planning a Route (AMS). 3. Sensing and building a Representations of the external Environment (ESS). 4. Exchanging such Representations and other Data with other CAVs and CAV-aware entities (AMS). 5. Making decisions about how to execute the Route (AMS). 6. Acting on the MAS. |
| CAV | Data | The digital representation of information. |
| CAV | Data Format | The standard digital representation of Data. |
| CAV | Data Type | An elementary component of a Data Format. |
| CAV | Data Format Compatibility | The ability of a Data Format to be losslessly converted to another Data Format. |
| CAV | Decision Horizon | The time within which a decision is assumed will be implemented. |
| CAV | Descriptor | Coded representation of text, audio, speech, or visual feature. |
| CAV | Digital Representation | A data structure corresponding to a physical entity. |
| CAV | Identifier | A Label that is uniquely associated with CAV. |
| CAV | Reference Model | The collection of the following resources:   1. AIW and their input/output data. 2. AIMs and their input/output data and connections. |
| CAV | Scene Descriptors | The individual attributes of the coded representation of the objects in a scene, including their location. |
| CAV | Subsystem | One of the 4 components making up the CAV. |
| ESS | Audio Object | Digital Representation of Audio information with its metadata. |
| ESS | Basic Environment Representation (BER) | A Digital Representation of the Environment produced by the set of an Environment Sensing Technology in a CAV and an Offline Map or provided by another CAV in range. |
| ESS | Deliberative Action | An action resulting from the processing of data |
| ESS | Environment | The portion of the real world of current interest to the CAV. |
| ESS | Environment Sensing Technology (EST) | One of the technologies used to sense the Environment by the Environment Sensing Subsystem, including Offline Map. |
| ESS | Global Navigation Satellite System (GNSS) | One of the systems providing global navigation informationsuch as GPS, Galileo, Glonass, BeiDou, Quasi Zenith Satellite System (QZSS) and Indian Regional Navigation Satellite System (IRNSS). |
| ESS | Inertial Measurement Unit | An inertial positioning device, e.g., odometer, accelerometer, speedometer, gyroscope etc. |
| ESS | Offline Map | A previously created digital map of an Environment and associated metadata. |
| ESS | Object | A Representation of an object. |
| ESS | Scene Description | The organised collection of Descriptors enabling an object-based description of a scene. |
| ESS | Shape | The digital representation of the space occupied by a CAV. |
| ESS | Spatial Attitude | The set of 18 values (x,y,z,α,β,γ; their 1st order derivatives; and their 2nd order derivatives) corresponding to the Position (x,y,z) and Orientation (α,β,γ) representing an Object’s pitch, yaw and roll and their derivatives. If the object is rigid, the position of the object corresponds to that of a representative point of the object. |
| ESS | Visual object | Coded representation of Visual information with its metadata. |
| ESS | Environment Representation | A digital representation of the Environment produced by an Environment Sensing Technology in CAV. |
| HCI | Audio | Digital representation of an analogue audio signal sampled at a frequency between 8-192 kHz with a number of bits/sample between 8 and 32, and non-linear and linear quantisation. |
| HCI | Audio Object | Coded representation of Audio information with its metadata. An Audio Object can be a combination of Audio Objects. |
| HCI | Audio Scene | The Audio Objects of an Environment with location metadata. |
| HCI | Audio-Visual Object | Coded representation of Audio-Visual information with its metadata. An Audio-Visual Object can be a combination of Audio-Visual Objects. |
| HCI | Audio-Visual Scene | (AV Scene) The Audio-Visual Objects of an Environment with Object location metadata. |
| HCI | Avatar | An animated 3D object representing a real or fictitious person in a virtual space rendered to a physical space. |
| HCI | Avatar Model | An inanimate avatar exposing handles for animation. |
| HCI | Cognitive State | An element of the internal status of a human or avatar reflecting their understanding of the Environment, such as “Confused” or “Dubious” or “Convinced”. |
| HCI | Colour (of Speech) | The timber of an identifiable voice independent of a current Personal Status and language. |
| HCI | Command | High-level instructions whose execution used to move a CAV to an intended Pose. |
| HCI | Descriptor | Coded representation of audio, speech, or visual feature. |
| HCI | Emotion | An element of the internal status of a human or avatar resulting from their interaction with the Environment or subsets of it, such as “Angry”, and “Sad”. |
| HCI | Environment | A physical or virtual space. |
| HCI | Face | The portion of a 2D or 3D digital representation corresponding to the face of a human. |
| HCI | Identity | The label uniquely associated with a human, an avatar or an object. |
| HCI | Manifestation | The manner of showing the Personal Status, or a subset of it, in any one of Speech, Face, and Physical Gesture. |
| HCI | Meaning (Gesture) | Information extracted from the head, arms, hands, and fingers conveying a coded message. |
| HCI | Meaning (Text) | Information extracted from an input text such as syntactic and semantic information. |
| HCI | Modality | One of Text, Speech, Face, or Gesture. |
| HCI | Personal Status | The ensemble of information internal to a person, including Emotion, Cognitive State, and Attitude. |
| HCI | Personal Status (Face) | The Personal Status, or a subset of it, conveyed by Face. |
| HCI | Personal Status (Gesture) | The Personal Status, or a subset of it, conveyed by Gesture. |
| HCI | Personal Status (Speech) | The Personal Status, or a subset of it, conveyed by Speech. |
| HCI | Physical Gesture | A movement of the body or part of it, such as the head, arm, hand, and finger, often a complement to a vocal utterance. |
| HCI | Response | Feedback autonomously generated by a CAV Subsystem in response to a Command. |
| HCI | Scene Description | The digital representation of an Environment populated by humans and real objects. |
| HCI | Social Attitude | An element of the internal status of a human or avatar related to the way they intends to position themselves vis-à-vis the Environment or subsets of it, e.g., “Confrontational”, “Respectful”. |
| HCI | Speech | Digital representation of analogue speech sampled at a frequency between 8 kHz and 96 kHz with 8, 16 and 24 bits/sample, and non-linear and linear quantisation. |
| HCI | Text | A series of characters drawn from a finite alphabet. |
| HCI | Viewpoint | The Spatial Attitude of a user looking at the Environment. |
| HCI | Visual Object | Coded representation of Visual information with its metadata. A Video Object can be a combination of Video Objects. |
| HCI | Vocal Gesture | Utterance, such as cough, laugh, hesitation, etc. Lexical elements are excluded. |
| HCI | Word Lattice | A directed graph containing speech recognition word candidates. |

# References

## Normative References

This document references the following normative documents:

1. Technical Specification: Governance of the MPAI Ecosystem (MPAI-GME) V1.1; https://mpai.community/standards/mpai-gme/.
2. Technical Specification: AI Framework (MPAI-AIF) V1.1; https://mpai.community/standards/mpai-aif/.
3. Technical Specification: Avatar Representation and Animation (MPAI-ARA), WD0.7; https://mpai.community/standards/mpai-ara/.
4. Multimodal Conversation (MPAI-MMC) V1.2; https://mpai.community/standards/mpai-mmc/.
5. Technical Specification: Context-based Audio Enhancement (MPAI-CAE) V2, https://mpai.community/standards/mpai-cae/.
6. Universal Coded Character Set (UCS): ISO/IEC 10646; December 2020
7. ISO/IEC 14496-10; Information technology – Coding of audio-visual objects – Part 10: Advanced Video Coding.
8. ISO/IEC 23008-2; Information technology – High efficiency coding and media delivery in heterogeneous environments – Part 2: High Efficiency Video Coding.
9. ISO/IEC 23094-1; Information technology – General video coding – Part 1: Essential Video Coding.

# Human-CAV Interaction (HCI)

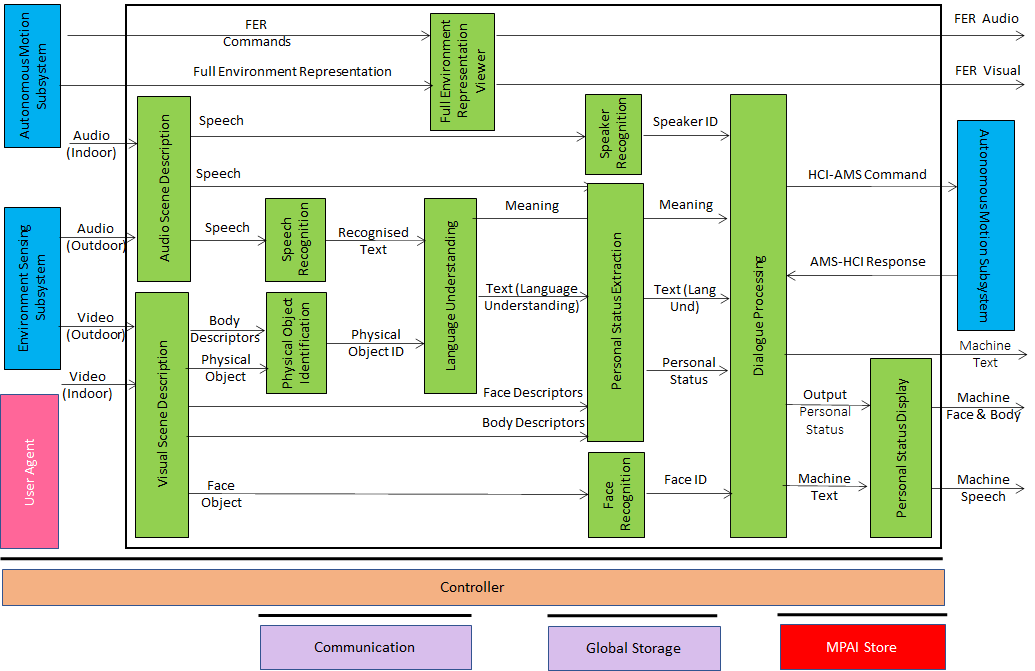
## Subsystem description

The Human-CAV Interaction (HCI) Subsystem performs the following high-level functions:

1. Authenticates humans e.g., for the purpose of letting them into the CAV.
2. Interprets and executes commands provided by humans e.g., to go to a Waypoint, display Full Environment Representation to passengers, turn off air conditioning, open window, or door, call a person, search for information etc.
3. Interprets conversation utterances with the support of the extracted Personal Statuses of the humans, e.g., on the fastest way to reach a Waypoint because of an emergency, or during a casual conversation.
4. Appears as a Body and Face with a mouth uttering Speech, Body, Face, and Speech show a Personal Status comparable to the Personal Status that a human driver would display in similar circumstances.

## Reference architecture

*Figure 4* gives the Human-CAV Interaction (HCI) Reference Model supporting the case of a group of humans approaching the CAV from outside the CAV and sitting inside the CAV.



*Figure 4 – Human-CAV Interaction Reference Model*

1. When a group of humans approaches the CAV from outside the CAV:
   1. The Audio Scene Description AIM creates the Audio Scene Descriptions in the form of Audio (Speech) Objects corresponding to each speaking human in the Environment (close to the CAV).
   2. The Visual Scene Description creates the Visual Scene Description and provides 1) the Face and Physical Objects and 2) the Body and Face Descriptors corresponding to each human in the Environment (close to the CAV).
   3. The Speaker Recognition and Face Recognition AIMs authenticate the humans the HCI is interacting with.
   4. The Speech Recognition AIM recognises the speech of each human.
   5. The Language Understanding AIM produces the refined Text (Language Understanding) and extracts the Meaning.
   6. The Personal Status Extraction AIM extracts the Personal Status of the humans from 1) Speech, 2) Face and Body Descriptors, 3) Text (Language Understanding) and 4) Meaning.
   7. The Dialogue Processing AIM 1) validates the human Identities, 2) responds to human utterances, 3) displays the Face and Body of the HCI Personal Status, and 4) issues commands to the Autonomous Motion Subsystem.
2. When a group of humans sit inside the CAV:
   1. The Audio Scene Description AIM creates the Audio Scene Descriptions in the form of Audio (Speech) Objects corresponding to each speaking human in the cabin.
   2. The Visual Scene Description creates the Visual Scene Descriptors in the form of Face Descriptors corresponding to each human in the cabin.
   3. The Speaker Recognition and Face Recognition AIMs identify the humans the HCI is interacting with.
   4. The Speech Recognition AIM recognises the speech of each human.
   5. The Language Understanding AIM extracts the Meaning and produces the refined Text (Language Understanding).
   6. The Personal Status Extraction AIM extracts the Personal Status of the humans.
   7. The Dialogue Processing AIM validates the human Identities, responds to human utterances, displays the HCI Personal Status, and issues commands to the Autonomous Motion Subsystem.
3. The HCI interacts with the humans sitting in the cabin in two ways:
   1. By responding to commands/queries from one or more humans at the same time, e.g.:
      1. Commands to go to or park at a Waypoint, etc.
      2. Commands with an effect on the cabin, e.g., turn off air conditioning, turn on the radio, call a person, open window or door, search for information etc.

Note: For completeness, *Figure 4* includes the conversion of human commands and responses from the CAV. However, this document does not address the format in which the HCI interacts with the Autonomous Motion Subsystem.

* 1. By conversing with and responding to questions from one or more humans at the same time about travel-related issues (in-depth domain-specific conversation), e.g.:
     1. Humans request information, e.g., time to destination, route conditions, weather at destination, etc.
     2. Humans ask questions about objects in the cabin or held by humans.
     3. CAV offers alternatives to humans, e.g., long but safe way, short but likely to have interruptions.
  2. By following the conversation on travel matters held by humans in the cabin if:
     1. The passengers allow the HCI to do so, and
     2. The processing is carried out inside the CAV.

1. While in the cabin, passengers can become aware of the external Environment by navigating the Full Environment Representation.
2. When conversing with the humans in the cabin, the HCI manifests itself as a speaking avatar via the Personal Status Display.

## I/O Data of the HCI AIW

*Table 3* gives the input/output data of Human-CAV Interaction.

*Table 3 – I/O data of* *Human-CAV Interaction*

|  |  |  |
| --- | --- | --- |
| **Input data** | **From** | **Comment** |
| Audio (ESS) | Environment Sensing Subsystem | User authentication  User command  User conversation |
| Audio | Cabin Passengers | User’s social life  Commands/interaction with HCI |
| Video (ESS) | Environment Sensing Subsystem | Commands/interaction with HCI |
| Video | Cabin Passengers | User’s social life  Commands/interaction with HCI |
| Full Environment Representation | Autonomous Motion Subsystem | Rendered by Full Environment Representation Viewers |
| Full Environment Representation Commands | Cabin Passengers | To control rendering of Full Environment Representation |
| **Output data** | **To** | **Comments** |
| Output Speech | Humans in Environment  Cabin Passengers | HCI’s response to passengers |
| Output Face | Cabin Passengers | HCI’s face when speaking |
| Output Body | Cabin Passengers | HCI’s body when speaking |
| Output Text | Cabin Passengers | HCI’s response to passengers |
| Full Environment Representation Audio | Passenger Cabin | For passengers to hear external Environment |
| Full Environment Representation Video | Passenger Cabin | For passengers to view external Environment |

## I/O Data of the HCI AIMs

*Table 4* gives the functions of the Human-CAV Interaction AIMs.

*Table 4* *– AI Modules of* *Human-CAV Interaction*

|  |  |  |
| --- | --- | --- |
| **AIM** | **Input** | **Output** |
| **Audio Scene Description** | Environment Audio (outdoor)  Environment Audio (indoor) | Speech Objects |
| **Visual Scene Description** | Environment Video (outdoor)  Environment Video (indoor) | Face Objects  Physical Objects  Body Descriptors  Face Descriptors |
| **Speech Recognition** | Speech Object | Recognised Text |
| **Physical Object Identification** | Physical Object  Human Object | Object ID |
| **Full Environment Representation Viewer** | FER Commands | FER Audio  FER Visual |
| **Language Understanding** | Recognised Text  Personal Status  Object ID | Meaning  Personal Status  Text (Language Understanding) |
| **Speaker Recognition** | Speech Descriptors | Speaker ID |
| **Personal Status Extraction** | Recognised Text  Speech Object  Face Object  Human Object | Personal Status |
| **Face Recognition** | Face Object | Face ID |
| **Dialogue Processing** | Speaker ID  Meaning  Text (Language Understanding)  Personal Status  Face ID  AMS-HCI Response | AMS-HCI Commands  Output Text  Output Personal Status |
| **Personal Status Display** | Machine Text  Output Personal Status | Machine Face & Body  Machine Text  Machine Speech |

## Data Formats

### HCI-AMS Commands

#### Functionality

The Dialogue Processing AIM

1. Understands and converts the commands that a User issues to the HCI to a form that the Autonomous Motion Subsystem can understand and execute.
2. Forwards messages that the Ego HCI Subsystem has received from the HCI Subsystems of other CAVs.

#### Requirements

Possible Commands are:

1. Provide Route options connecting two Poses at a given time.
2. Execute Route
3. Suspend Route
4. Resume Route

#### Technology

A language enabling the HCI Subsystem to communicate Commands or forward information from other HCI Subsystems to the Autonomous Motion Subsystem.

### AMS-HCI Response

#### Functionality

The Autonomous Motion Subsystem can

1. Respond to a Command.
2. Provide HCI with information (e.g., about unexpected stops).

#### Requirements

Possible Responses are:

1. List of Road options
2. Unexpected delay
3. Road problems
4. Equipment failure
5. Target reached.

#### Technology

A language enabling the Autonomous Motion Subsystem to communicate Respond to Commands or information from other HCI Subsystems to the HCI Subsystem.

### Audio

#### Functionality

A serialised output of microphones or a microphone array is used to capture the external sound.

#### Requirements

Audio Data shall be usable for the following purposes:

1. Create the Audio Scene Description to:
   1. Enable extraction of speech addressed by humans outdoor or indoor to the HCI.
   2. Incorporate outdoor Audio information into the Basic Environment Representation.
2. Suppress noise and individual sound sources outside and inside the passenger cabin.

#### Technology

MPAI-CAE V2 [5] specifies Interleaved Multichannel Audio for use in an outdoor and indoor environment.

### Audio Scene Descriptors

#### Functionality

The Audio Scene Descriptors shall enable the description of the outdoor and indoor sound field in terms of Audio Objects.

#### Requirements

The Audio Scene Descriptors shall:

1. Identify individual Audio Objects.
2. Provide the Spatial Attitude of each Audio Object.
3. The Spatial Attitude of each Audio Object can be represented as Cartesian coordinates of the source and their time derivatives.

#### Technology

MPAI-CAE V2 [5] specifies Audio Scene Descriptors as φ (Azimuth) and θ (Elevation) of the sound sources of each Audio Object.

### Visual Scene Descriptors

#### Functionality

The Visual Scene Descriptors shall enable the description of the outdoor and indoor visual scene in terms of Visual Objects.

#### Requirements

The Audio-Visual Scene Descriptors shall:

1. Identify individual Visual Objects.
2. Provide the Spatial Attitude of each Visual Object.
3. The Spatial Attitude of each Visual Object shall be represented as Cartesian coordinates of the Visual Object and their time derivatives.

#### Technology

TBD

### Visual Scene Descriptors

#### Functionality

The Audio-Visual Scene Descriptors shall enable the description of the outdoor and indoor audio-visual scene.

#### Requirements

All Audio-Visual Scene Descriptors shall:

1. Have a common time base.
2. Enable the association of co-located audio and visual objects if both are available.
3. Support the physical displacement and interrelation (e.g., occlusion) of audio and visual objects over time.

For example, the description of the audio-visual objects in the Real Environment where a group of speaking humans with a background noise.

#### Technology

TBD

### Avatar Model

#### Functionality

A user can select the Avatar Model used by the CAV to manifest itself though a Personal Status Display AIM.

#### Requirements

The Avatar Model shall operate within the constraints of the Personal Status Display.

#### Technology

MPAI is developing Avatar Representation and Animation Technical Specification [3]that includes an Avatar Model with the given requirements.

### Face Descriptors

#### Functionality

The Face Descriptors are used:

1. To extract the Personal Status of a face.
2. To identify a human from their face.
3. To animate the Face of an Avatar Model.
4. To express the Personal Status of Face in an Avatar Model.

#### Requirements

The Face Descriptors shall:

1. Represent time-dependent standard Cognitive State, Emotion, and Social Attitude values in the Face of a standard Avatar Model.

#### Technology

MPAI is developing Avatar Representation and Animation Technical Specification [3] that includes Face Descriptors supporting the requirements.

### Face ID

#### Functionality

Face ID is the Identifier of the human.

#### Requirements

Face identification should support the members of a restricted group (e.g., a family) or the customers of a car renting company.

#### Technology

TBD.

### Face Object

#### Functionality

A Face Object is the 2D or 3D digital representation of a face or of an Avatar Face.

#### Requirements

It shall be possible to extract a Face Object from a Visual Scene Description.

#### Technology

TBD

### Full Environment Representation Commands

#### Functionality

A CAV passenger issues FER Commands to navigate the Full Environment Representation:

#### Requirements

The possible commands are:

1. Select a Point of View.
2. Zoom in/out.
3. Control sound level.

#### Technology

TBD

### Full Environment Representation Audio

#### Functionality

The audio output of the Full Environment Representation Viewer.

#### Requirements

FER Audio should be compatible with commercially available headsets and Head Mounted Displays.

#### Technology

TBD

### Full Environment Representation Visual

#### Functionality

The visual output of the Full Environment Representation Viewer.

#### Requirements

FER Visual should be compatible with commercially available displays and Head Mounted Displays.

#### Technology

TBD

### Full Environment Representation

#### Functionality

The FER results from the fusion of the Basic Environment Representation (BER) developed by the Ego Environment Sensing Subsystem and the versions of the BERs developed by other CAVs.

#### Requirements

The requirements of the Full Environment Representation data format are developed in the context of the Autonomous Motion Subsystem.

#### Technology

See Section 7.4.7 – Autonomous Motion Subsystem.

### Body Descriptors

#### Functionality

The Body Descriptors are used:

1. To extract the motion of a body.
2. To extract the Personal Status of a body.
3. To animate an Avatar Model.
4. To express the Gesture in an Avatar Model.

#### Requirements

The Body Descriptors shall represent:

1. The motion of a body.
2. The time-dependent standard Cognitive State, Emotion, and Social Attitude values of a standard Avatar Model.

#### Technology

MPAI is developing Avatar Representation and Animation Technical Specification [3] that includes Body Descriptors supporting the requirements.

### Machine Face & Body

The rendering of the face and body of an avatar representing the HCI Subsystem.

### Machine Speech

The rendering of the synthetic speech generated by the HCI.

### Meaning

#### Functionality

Meaning is the result of natural language analysis

#### Requirements

Meaning shall include the following elements:

* POS\_tagging
* NE\_tagging
* Dependency\_tagging
* SRL\_tagging

#### Technology

MPAI has developed the Multimodal Conversation Technical Specification (MPAI-MMC) [4] specifying Meaning.

### Physical Object ID

#### Functionality

The Identifier uniquely associates with a Physical Object.

#### Requirements

TBD

#### Technology

MPAI has developed the Multimodal Conversation Technical Specification (MPAI-MMC) [4] specifying ObjectID.

### Orientation

#### Functionality

Orientation expresses the direction of an Object.

#### Requirements

Orientation is defined based on…

#### Technology

### Personal Status

#### Functionality

Personal Status (PS) indicates a set of 3 Factors {Cognitive State, Emotion, Social Attitude} conveyed by one or more of the Text, Speech, Face, and Gesture Modalities.

#### Requirements

See

#### Technology

### Physical Objects

#### Functionality

#### Requirements

An AIM shall be able to extract a Physical Object present in a Visual Scene Description.

#### Technology

### Position

##### Functionality

The Position of an Object refers to the coordinates of a representative point of an object in a coordinate system.

##### Requirements

The coordinates can be:

1. On the surface of the Earth, e.g., in the case of the objects in a Environment Representation.
2. Defined with reference to a specific point on the Earth.
3. Possibly arbitrarily defined local coordinates. The specific use case defines the point with coordinates (0,0,0).
4. In a virtual space. The specific use case defines the point with coordinates (0,0,0).

##### Technology

### Speaker ID

##### Functionality

##### Requirements

##### Technology

### Speech Descriptors

##### Functionality

Speech Descriptors are used to recognise the identity of a limited number of humans, e.g.,

1. Members of a family.
2. Customers of a CAV-renting company.

##### Requirements

##### Technology

### Speech Object

##### Functionality

##### Requirements

An AIM shall be able to extract a Speech Object present in a Visual Scene Description.

##### Technology

### Text

##### Functionality

A Character Set shall be able to represent the characters used by most currently used languages.

##### Requirements

In MPAI-MMC V2 Recognised Text shall be represented in a format that provides for each selected unit, e.g., word, syllable, phoneme:

1. Timestamped start.
2. Duration
3. Strings of characters or phonemic symbols
4. Corresponding probability.

##### Technology

All the following instances of Text:

1. Text (Language Understanding)
2. Machine Text
3. Output Text

shall be represented by [4].

Recognised Text may also be represented by [4]

### Video

##### Functionality

##### Requirements

The visual field can be captured by one or an array of visual sensors with or without depth information. The output can have different resolutions, colour spaces, and frame rates.

##### Technology

### Visual Scene Descriptors

##### Functionality

Description of a visual scene where humans are:

1. Outdoor, approach a CAV and are recognised by it.
2. Inside a CAV and converse with it.

##### Requirements

Visual Scene Descriptors describe the structured composition of Visual Objects in a Visual Scene. The Format shall enable an application to extract an individual object from a scene.

##### Technology

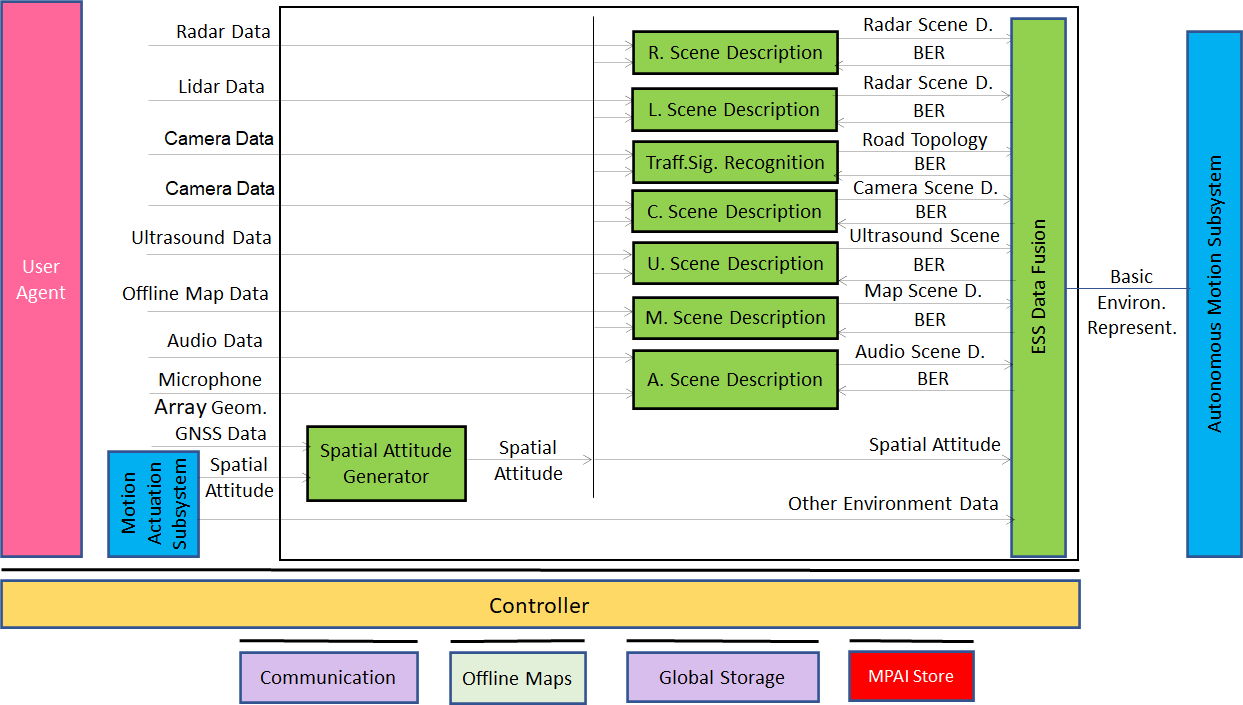
# Environment Sensing Subsystem (ESS)

## Subsystem description

The purpose of the Environment Sensing Subsystem (ESS) is to acquire as much as possible information from the Environment – electromagnetic, acoustic, spatial, and other physical (e.g., temperature, pressure, humidity etc.) – using the devices onboard the CAV with the goal of creating the Basic Environment Representation (BER).

## Reference architecture

*Figure 5* gives the Environment Sensing Subsystem Reference Model.



*Figure 5 – Environment Sensing Subsystem Reference Model*

The ultimate purpose of the ESS is to produce the Basic Environment Representation (BER) and the Spatial Attitude of the Ego CAV using all Sensed Data from available ESTs, the Offline Maps, and the information received from Roadside Units.

*Figure 5* represents the general case in which snapshots of raw information provided by an Environment Sensing Technology (EST), e.g., Radar Data, is acquired. As such the information acquired may not be in a form suitable for the subsequent high-level data processing. An EST-specific Data Processor, e.g., the Radar Data Processor AIM is invoked. The output data are then passed, e.g., to the Radar Scene Description AIM which produces the Radar Scene Descriptors. The model of *Figure 5* assumes that there are two AIMs one processing the acquired data (e.g., the Radar Data Processor AIM) and one using the processed data (e.g., the Radar Scene Description AIM) to create EST-specific Scene Descriptors. However, a specific implementation may integrate the two AIMs into a single AIM.

An EST-specific Scene Description (SD) is a structured combination of Scene Descriptors with the following features:

1. It may be a complex data structure that includes several elementary components called Data Types each having a Data Format.
2. It is object-based, time-dependent, and constantly updated.
3. Objects that may have different resolutions, e.g., the background may be a single object with low resolution while a CAV that is 5 m away from the Ego CAV has a high resolution.
4. It is compatible with the SD obtained from a different EST, i.e., the SD1 from EST1 can be converted to the SD2 of EST2 without loss of information.
5. It may include Data Types not included in the SD of another EST and vice versa.
6. Data in SD#1 of EST#1 need not have the same values even though the Data Types may be the same because their values may:
   1. Not have the same Accuracy.
   2. Conflict. even though their Data Types may be the same or compatible.

The Online Map Data is treated as an EST. The formats of the Offline Maps shall allow for integration in an SD without loss of information.

*Figure 5* (and the rest of this Chapter) assumed that the Traffic Signalisation Recognition produces the Road Topology by analysing the Processed Camera Data. The model can easily by extended to the case where another EST is processed to produce the Road Topology.

## Input and output data

The Environment Sensing Technologies (EST) currently considered by MPAI-CAV are:

1. Global navigation satellite system or GNSS (~1 & 1.5 GHz Radio).
2. Geographical position and orientation, and their time derivatives up to 2nd order.
3. Data in the visible range, possibly supplemented by depth information (400 to 700 THz).
4. Lidar (~200 THz infrared).
5. Radar (~25 & 75 GHz).
6. Ultrasound (> 20 kHz).
7. Audio in the audible range (16 Hz to 16 kHz).
8. Spatial Attitude (in addition to GNSS, from the Motion Actuation Subsystem).
9. Other environmental data (temperature, humidity, ...).

*Table 5* gives the input/output data of Environment Sensing Subsystem.

*Table 5 – I/O data of* *Environment Sensing Subsystem*

|  |  |  |
| --- | --- | --- |
| **Input data** | **From** | **Comment** |
| Spatial Attitude from AMS | Motion Actuation Subsystem | To be fused with GNSS data |
| Other Environment Data | Motion Actuation Subsystem | Temperature etc. to be added to Basic Environment Representation |
| Global Navigation Satellite System (GNSS) | ~1 & 1.5 GHz Radio | Get Pose from GNSS |
| Radar | ~25 & 75 GHz Radio | Capture Environment with Radar |
| Lidar | ~200 THz infrared | Capture Environment with Lidar |
| Ultrasound | Audio (>20 kHz) | Capture Environment with Ultrasound |
| Cameras (2/D and 3D) | Video (400-800 THz) | Capture Environment with Cameras |
| Microphones | Audio (16 Hz-16 kHz) | Capture Environment with Microphones |
| **Output data** | **To** | **Comment** |
| Alert | Autonomous Motion Subsystem | Critical last minute Environment Description from EST |
| Basic Environment Representation | Autonomous Motion Subsystem | Locate CAV in the Environment |

## AI Modules

*Table 6* gives the functionality of all Environment Sensing Subsystem AIMs.

*Table 6 – AI Modules of* the *Environment Sensing Subsystem*

|  |  |
| --- | --- |
| **AIM** | **Function** |
| **Ultrasound Data Processor** | Processes ultrasound data. |
| **Audio Data Processor** | Processes microphone array data. |
| **Online Map Data Processor** | Converts online map data to a format enabling the creation of the online map-based Scene Description. |
| **Radar Scene Description** | Produces Radar Scene Descriptors from Processed Radar Data |
| **Lidar Scene Description** | Produces Lidar Scene Descriptors from Processed LidarData |
| **Traffic Signalisation Recognition** | Produces Road Topology of the Environment. |
| **Ultrasound Scene Description** | Produces Ultrasound Scene Descriptors from Processed Ultrasound Data. |
| **Audio Scene Description** | Produces Audio Scene Descriptors from Processed Audio Data. |
| **Online Map Scene Description** | Produces Online Map Data Scene Descriptors from Processed Online Map Data. |
| **Environment Sensing Subsystem Data Fusion** | Selects critical Environment Representation as Alert; produces CAV’s Basic Environment Representation by fusing the Scene Descriptors of the different ESTs, Road Topology, Spatial Attitude, and Other Environment Data. |

## I/O Data summary

For each AIM (1st column), *Table 7* gives the input (2nd column) and the output data (3rd column). The following 3-digit subsections give the requirements of the data formats in columns 2 and 3.

*Table 7 – Environment Sensing Subsystem AIMs and data*

|  |  |  |
| --- | --- | --- |
| **AIM** | **Input** | **Output** |
| **Audio Data Processor** | Microphone Array Data. | Processed Audio Data. |
| **Radar Scene Description** | Radar Data.  Basic Environment Representation. | Radar Scene Descriptors. |
| **Lidar Scene Description** | Lidar Data.  Basic Environment Representation. | Lidar Scene Descriptors. |
| **Traffic Signalisation Recognition** | Camera Data.  Basic Environment Representation. | Road Topology. |
| **Camera Scene Description** | Camera Data.  Basic Environment Representation. | Lidar Scene Descriptors. |
| **Ultrasound Scene Description** | Ultrasound Data.  Basic Environment Representation. | Ultrasound Scene Descriptors. |
| **Audio Scene Description** | Audio Data  Basic Environment Representation. | Audio Scene Descriptors. |
| **Spatial Attitude Generation** | GNSS Data  Spatial Attitude form MAS | Spatial Attitude |
| **Environment Sensing Subsystem Data Fusion** | Radar Scene Descriptors.  Lidar Scene Descriptors.  Road Topology.  Lidar Scene Descriptors.  Ultrasound Scene Descriptors.  Audio Scene Descriptors.  Map Scene Descriptors.  Spatial Attitude.  Other Environment Data, | Basic Environment Representation.  Alert. |

## Data Formats

### Radar Data

RADAR Data is serialised data provided by a RADAR sensor.

#### Functionality

A RADAR is an active time-of-flight sensor operating in the 08-10 cm range.

#### Requirements

The RADAR Data Format shall include the following information:

1. RADAR frequency used.
2. Frame frequency.
3. Sample rate.
4. Bits per spatial coordinate and distance.
5. Multiplexed RADAR Data Format.

#### Technology

### LiDAR Data

LiDAR Data is serialised data provided by a LiDAR sensor.

#### Functionality

A LiDAR is an active time-of-flight[[2]](#footnote-2) sensor operating in the µm range – ultraviolet, visible, or near infrared light.

#### Requirements

The LiDAR Data Format shall include the following information:

1. Number and values of the LiDAR frequencies used.
2. Frame frequency. Frame duration is the time it takes for the LiDAR to make a full 360º scan.
3. Sample rate.
4. Bits per (x,y,x) coordinates.
5. Multiplexed Lidar Data Format.

#### Technology

### Camera Data

Camera Data is serialised data provided by a variety of visual sensor configurations.

#### Functionality

Camera Data can be provided by:

1. One video camera providing RGB (Red-Green-Blue) data.
2. An array of video cameras each providing RGB data.
3. One camera providing RGBD (RGB+Depth) data.
4. Av array of cameras ech providing RGBD (RGB+Depth) data.

#### Requirements

The Camera Data Format shall include the following information:

1. Camera geometry, including Depth.
2. Colour space.
3. Frame frequency.
4. Number of bits/sample.
5. Format of Multiplexed RGBD data.
6. Multiplexed RGBD data.

#### Technology

### Ultrasound Data

Ultrasound Data is serialised data provided by an ultrasonic sensor measuring the distance between objects within close range.

#### Functionality

An Ultrasonic sensor is an active time-of-flight sensor typically operating in the 40 kHz and 250 kHz range.

#### Requirements

#### Technology

The Ultrasound File Format initiative has defined the Ultrasound File Format (UFF) format [30].

### Offline Map Data

The Offline Map is a roadmap with centimetre-level accuracy and a high environmental fidelity.

#### Functionality

For a given set of coordinate values, the Offline Map provides the position of and information about:

1. Pedestrian crossings.
2. Traffic lights.
3. Traffic signs.
4. Barriers.
5. Sidewalks.
6. Grade of the road.
7. Sped limits.
8. Landmarks names.

at the time the Offline Map has been created.

#### Requirements

#### Technology

### Audio Data

#### Functionality

The serialised output of a microphone array capturing the target Environment.

#### Requirements

Audio Data shall be usable for the following purposes:

1. Create the Audio Scene Description (ASD) to:
   1. Enable extraction of speech addressed to CAV.
   2. Incorporate the ASD in the Basic Environment Representation.
2. Suppress noise and cancel individual sound source outside and inside the passenger cabin.

#### Technology

MPAI-CAE V1 [5] specifies Interleaved Multichannel Audio for use in a room environment which can be used for thr in-cabin case. MPAI is developing MPAI-CAE V2 will cover the case of an outdoor Environment when humans approach the CAV.

### Microphone Array Geometry

#### Functionality

Audio Data is captured by an array of microphones. The disposition of the microphones in the array is specified by the Microphone Array Geometry.

#### Requirements

Microphone Array Geometry data structure shall include:

1. Sensing characteristics of the microphone(s) used (e.g., cardioid),
2. Microphone array geometry.
3. Sampling frequency.
4. Number of bits/sample.

#### Technology

MPAI-CAE V1 [5] specifies a Microphone Array Geometry usable for the in-cabin use cases. MPAI-CAE V2 t will cover the case of a geometry suitable for the outdoor Environment when humans approach the CAV.

### GNSS Data

Global Navigation Satellite Systems (GNSS) provide spatial information with different accuracies.

#### Functionality

GNSS can only be relied on when reception conditions are above a certain level. Tunnels or urban canyons are unsuitable for GNSS use.

#### Requirements

GNSS data shall express Latitude and Longitude of a given point in degrees, minutes, seconds, and thousandths of second.

#### Technology

### Spatial Attitude

#### Functionality

The CAV Spatial Attitude is received from GNSS, MAS, and Roadside Units. It is measured with respect to:

1. A predetermined point in the CAV that is the origin and has (0,0,0) coordinates and (x,y) Cartesian coordinates.
2. The rotation angle φ with respect to a given direction.
3. The elevation angle θ with respect to the x,y plane.

|  |  |
| --- | --- |
|  | Roll-pitch-yaw angles of cars and other land based vehicles ... |

Figure 6 – Spatial Attitude in a CAV

#### Requirements

The Spatial Attitude shall be represented by the following:

1. Timestamp.
2. Longitude (x) and Latitude (y) expressed in degrees, minutes, seconds, and thousandths of second and Height (z) expressed in metres. The Environment is assumed to be sufficiently small (~100 m) that the ego CAV:
   1. Need not consider the effect of the curvature of the earth. The x,y plane is perpendicular to the line across the position of the CAV and the centre of the earth.
   2. Its Height is 0 and all other objects, including CAVs have a Height measured with respect to the ego CAV zero Height.
3. The angles of the ego CAV with:
   1. The x axis in the x,y plane measured in radians.
   2. The z axis measured in radians.
4. First and second order derivatives of x,y,φ,θ.

#### Technology

### Other Environment Data

In general, the CAV may have additional sensors located in the Motion Actuation Subsystem or receive relevant information online.

#### Functionality

The types of Environment Data provided by the MAS are:

1. Outside temperature.
2. Outside humidity.
3. Wind.
4. Air pressure.

#### Requirements

Information is expressed as:

1. Temperature: ºC.
2. Wind
   1. Velocity: m/s
   2. Direction: in degrees, minutes, seconds measured from the current meridian pointing to North rotating clockwise.
3. Air pressure: N/m2
4. Humidity: in g/m3

#### Technology

### Scene Description Format

#### Functionality

An EST-specific Scene Description (SD) format is a list of Scene Descriptors in the EST-specific time window.

#### Requirements

The SD at time t (SD(t)) shall contain all information that an EST-specific Scene Description AIM has available that can be usefully processed by a subsequent AIM, namely:

1. Timestamp
2. EST-specific Scene Description AIM ID.
3. Spatial Attitude of EST-specific Sensors.
4. List of Objects detected and confirmed at time t with their Metadata.

An EST-specific SD at time t SD(t) can be created with the following process:

1. For each Object in SD(t-Δt):
   1. Find the corresponding Object in the currently sensed Scene Description SDs(t).
   2. Update the attributes of that SDs(t) Object.
   3. Add that Object to SD(t).
   4. Do not add to SD(t) the Objects present in SD(t-Δt) and no longer present in SDs(t).
2. For each new Objects in SDs(t)
   1. Assign attributes to new Objects in SDs(t) (i.e., entirely new objects or the merge of two or more objects present in SD(t-Δt)).
   2. Add the new Objects to SD(t).

A Scene can have a 2D, 2.5D, or 3D Format that depends and its Objects can have different representations depending on whetherthey are static or dynamic.

**2D Scene**

1. Static Objects:
   1. Parametric free space representation represented as a single Object.
   2. Individual static Objects.
2. Dynamic Objects: object-based representation.

**2.5D Scene**

1. Static Objects:
   1. Grid-based (elevation maps or stixel world), represented as a single Object.
   2. Object-based for traffic poles and signals (e.g., Stixel world, Multi-level surface map).
2. Dynamic Object: Object-based (e.g., Stixel world, Multi-level surface map).

**3D Scene**

1. Static Objects
   1. Voxel grids, meshes, possibly as a single Object.
   2. Object-based for traffic poles and signals (voxel grids, meshes).
2. Dynamic Objects: point clouds, voxel grids, meshes, …

#### Technology

### Object Format

#### Functionality

An Object is a data structure representing an object sensed by an EST and produced by an EST-specific Scene Description.

#### Requirements

An Object shall contain the following elements:

1. Timestamp.
2. Identifier of the Scene Description AIM that has generated the Object.
3. Alert
4. Spatial Attitude of the Object and its estimated accuracy measured from the CAV Center.
5. Bounding box.
6. Object type:
   1. 2D Object Description.
   2. 2.5D Object Description
   3. 3D Object Description
7. State (Static, Dynamic, Unknown and their estimation accuracies)
8. Identifier of Object
   1. Already existing and mapping to the Object and estimated accuracy.
   2. Found to be new.
9. Any Semantics available at this stage.

#### Technology

### Lidar Scene Description

#### Functionality

The LiDAR Scene Description AIM

1. Receives LiDAR Data.
2. Requests portions of a previous Basic Environment Representation.
3. Provides LiDAR Scene Descriptors including any Alert.

The LiDAR sensor provides information about the cloud of points P(x,y,z) as (φ,θ,r). The x,y,z coordinates are obtained from the following formulas:

x=rsin(θ)cos(φ)

y=rcos(θ)cos(φ)

z=rsin(θ)

where x,y,z have the same definition used in Spatial Attitude.

#### Requirements

R is expressed in meters using N bits, and θ,φ in radian/degrees using N bits.

x,y,z are expressed with a precision of M bits.

The data stream has a structure TBD.

The Alert

#### Technology

### Radar Scene Description

#### Functionality

The RADAR Scene Description AIM

1. Receives RADAR Data.
2. Requests portions of a previous Basic Environment Representation.
3. Provides RADAR Scene Descriptors including any Alert.

#### Requirements

#### Technology

### Road Topology

#### Functionality

The Road Topology AIM

1. Receives Camera Data.
2. Requests portions of a previous Basic Environment Representation.
3. Provides Road Topology and Alert.

#### Requirements

The Road Topology includes:

1. The Position of the Road Signs (Traffic Poles, Road Signs, Traffic Lights).
2. The Taxonomy-based semantics of the Road Signs.

#### Technology

### Camera Scene Descriptors

#### Functionality

The Camera Scene Description AIM:

1. Receives Camera Data.
2. Requests portions of a previous Basic Environment Representation.
3. Provides Camera Scene Descriptors including any Alert

#### Requirements

#### Technology

### Ultrasound Scene Description

#### Functionality

The Ultrasound Scene Description AIM

1. Receives Ultrasound Data.
2. Requests portions of a previous Basic Environment Representation.
3. Provides Ultrasound Scene Descriptors including any Alert.

#### Requirements

#### Technology

### Maps Scene Description

#### Functionality

The Map Scene Description AIM

1. Receives
2. A coordinate point on the surface of the Earth.
3. Offline Map Data.
4. Requests portions of a previous Basic Environment Representation.
5. Provides Object-based Map Scene Descriptors.

#### Requirements

The Map Scene Description provides:

1. Object-based representation of the Environment around the input point.
2. Coordinates, size, and shape of the Objects.
3. Taxonomy-based semantics of the Objects.

#### Technology

### Audio Scene Description

#### Functionality

The Audio Scene Description AIM

1. Receives
2. Audio Data
3. Microphone Array Geometry
4. Requests portions of a previous Basic Environment Representation.
5. Provides Audio Scene Descriptors including any Alert.

#### Requirements

#### Technology

### Alert

#### Functionality

Alert is Data generated by a Scene Description AIM that signals a danger estimated by that AIM as information added to its Scene Description.

#### Requirements

Alert is represented by the Object that the Scene Description AIM has found worth attention.

#### Technology

### Basic Environment Representation

#### Functionality

The Basic Environment Representation (BER) results from the *integration* of the different Scene Descriptors generated by the different EST-specific Scene Description AIMs and Environmental data. It contains:

1. Audio and Visual Objects
2. Their Spatial Attitude
3. Scene Geometry
4. Environmental data (e.g., weather, temperature, air pressure, ice and water on the road, wind, fog etc.)
5. Any semantic information about the Scene.

#### Requirements

The functional requirements of the Basic Environment Representation (BER) are:

1. The BER shall include all available information from ESS amd MAS that enables CAV to define a Path in the Decision Horizon Time.
2. The BER shall include the BERs of snapshots sufficiently prior in time captured at a frequency useful to track object and their motion, infer motion vectors, etc.
3. The BER shall include:
   1. A description of the External Environment in terms of Road Topology (e.g., roads and lane geometry).
   2. The traffic rules and regulatory restrictions derived from Road Signs available at this stage and applying to relevant portions of the External Environment[[3]](#footnote-3).
4. The BER shall describe each Object with the following attributes:
   1. ID.
   2. Timestamp.
   3. Type (Audio or Visual).
   4. ID of related objects of a different Type.
   5. Motion State (Static, Dynamic, Unknown).
   6. Spatial Attitude.
   7. Alert flag
   8. Predicted Trajectory.
   9. Object dimensionality (2D, 2.5D and 3D). Applicable only to Visual Objects.
   10. Shape of Visual Objects
   11. Semantics (to the extent it is available at this stage) fo
       1. Another CAV or other object types
       2. An Object belonging to a group of Objects (listed by their IDs).
   12. Accuracy of all the estimated values in the Object.
5. The BER shall allow for easy check of the feasibility of a Trajectory (e.g., the AMS can easily check that the intended Trajectory of the ego CAV designed to reach the intended point does not collide with other Visual Objects in the Decision Horizon based on the current state of the BER).
6. The BER shall have a scalable representation, i.e., allowing for:
   1. Refinement of a BER using new EST-specific Scene Descriptors.
   2. Extraction of part of the BER based on a required Level of Detail (e.g., object bounding boxes and their Spatial Attitudes).
   3. Easy replacement of addition of new data (e.g., adding shape of an object when there was only the bounding box).
   4. Fast access to object metadata, e.g.:
      1. Spatial Attitude.
      2. Predicted Trajectory.
      3. Shape (e.g., bounding box for a Visual Object).
   5. Selected (read) access to data required by different AIMs, e.g.:
      1. Radar Scene Description accessing the current BER to improve its description.
   6. Possibility to update Object, Scene, or BER from one snapshot to another.
   7. Possibility that a CAV communicate to another CAV a subset of its BER containing Objects with different degrees of details, e.g., starting from bounding boxes and their Spatial Attributes, depending on the available bandwidth.

#### Technology

# Autonomous Motion Subsystem (AMS)

## Use Case description

The typical series of operations carried out by the Autonomous Motion Subsystem (AMS) is described below. Note that the sequential description does not imply that an operations can only be carried out after the preceding one has been completed.

1. Human-CAV Interaction requests Autonomous Motion Subsystem to plan and move the CAV to the human-selected Pose. AMS-human dialogue may follow.
2. Computes the Route satisfying the human’s request.
3. Receives the current Basic Environment Representation from Environment Sensing Subsystem.
4. While moving, CAV
   1. Broadcasts the Basic Environment Representation and other data.
   2. Receives Basic Environment Representations and other data from other CAVs.
   3. Produces the Full Environment Representation by fusing its own Basic Environment Representation with those from other CAVs in range.
   4. Plans a Path connecting Poses.
   5. Selects behaviour and motion to reach the next Pose acting on information about the Poses other CAVs in range intend to reach and the Objects between the current Pose and the next Pose.
   6. Defines a Trajectory that
      1. Complies with general traffic regulations and local traffic rules.
      2. Preserves passengers’ comfort.
   7. Refines Trajectory to avoid obstacles.
   8. Sends Commands to the Motion Actuation Subsystem to take the CAV to the next Pose.
5. Stores the data resulting from a decision (Route Planner, Path Planner etc.)

The AMS should be designed in such a way that different levels of autonomy, e.g., those indicated by SAE International [10], are possible depending on the amount and level of available func­tionalities.

## Reference architecture

*Figure 7* gives the Autonomous Motion Subsystem Reference Model.



*Figure 7 – Autonomous Motion Subsystem Reference Model*

This is the operation according to the Reference Model:

1. A human requests the Human-CAV Interaction to be transported to a destination.
2. This request is interpreted and passed to the AMS.
3. The AMS activates the Route Planner to generate a set of Waypoints starting from the current Pose, obtained from the Full Environment Representation, up to the destination.
4. The Waypoints enter the Path Planner which generates a set of Poses to reach the next Waypoint.
5. For each Path, the Motion Planner generates a Trajectory to reach the next Pose.
6. The Obstacle Avoider AIM receives the Tracjectory and checks if there is a last-minute change, detected from Alert.
7. If an Alert was received, the Obstacle Avoider AIM checks whether the implementation of the Trajectory createds a collision, especially with the Object creating the Alert.
   1. If a collision is indeed detected, the Obstacle Avoider AIM requests a new Trajectory from the Motion Planner.
   2. If no collision is detected, Obstacle Avoider AIM issues a Command to the Motion Actuation Subsystem.
8. The Motion Actuation Subsystem sends Feedback about the execution of the Command.
9. The decision of each element of the said chain may be recorded.

## Input and output data

*Table 8* gives the input/output data of Autonomous Motion Subsystem.

*Table 8 – I/O data of* *Autonomous Motion Subsystem*

|  |  |  |
| --- | --- | --- |
| **Input data** | **From** | **Comment** |
| Command from HCI | Human-CAV Interaction | Human commands, e.g., “take me home” |
| Basic Environment Representation | Environment Sensing Subsystem | CAV’s Environment representation. |
| Other V2X Data | Other CAVs | Other CAVs and vehicles, and roadside units. |
| Feedback from MAS | Motion Actuation Subsystem | CAV’s response to Command. |
| **Output data** | **To** | **Comment** |
| Response to HCI | Human-CAV Interaction | MAS’s response to AMS Command |
| Command to MAS | Motion Actuation Subsystem | Macro-instructions, e.g., “in 5s assume a given State”. |
| Full Environment Representation | Other CAVs | For information to other CAVs |

## AI Modules

*Table 9* gives the AI Modules of the Autonomous Motion Subsystem.

*Table 9 – AI Modules of Autonomous Motion Subsystem*

|  |  |
| --- | --- |
| **AIM** | **Function** |
| **Route Planner** | Computes a Route, through a road network, from the current to the target destination. |
| **Path Planner** | Generates a set of Paths, considering:   1. The Route. 2. Spatial Attitude. 3. Full Environment -Representation. 4. Traffic Rules. |
| **Motion Planner** | Defines a Goal and a Trajectory to reach the Goal using the Spatial Attitude satisfying the CAV’s kinematic and dynamic constraints, and considering passengers’ comfort. |
| **Obstacle Avoider** | Checks that the Trajectory is compatible with any Alert information. If it is, it passes the Trajectory to the Commnad Issuer. If it is not, it requests a new Trajectory. |
| **Command** | Instructs the MAS to execute the Trajectory considering the Environment conditions. |
| **Full Environment -Representation Fusion** | Creates an internal representation of the Environment by fusing infor­mation from itself, CAVs in range and other transmitting units. |

### Summary of Subsystem data

*Table 10* gives, for each AIM (1st column), the input data (2nd column) and the output data (3rd column).

*Table 10 – CAV Autonomous Motion Subsystem data*

|  |  |  |
| --- | --- | --- |
| **CAV/AIM** | **Input** | **Output** |
| **Route Planner** | Pose  Destination  Offline maps | Route  Estimated time |
| **Full Environment Representation** | Basic Environment Representations  Other V2X Data | Full Environment Representation |
| **Path Planner** | Offline maps  Route  Full Environment Representation | Set of Paths |
| **Motion planner** | Path  Full Environment Representation | Trajectory |
| **Obstacle Avoider** | Trajectory  Full Environment Representation | Trajectory  Route State |
| **Command to AMS** | Trajectory  Environment Data  Feedback | Command |
| **Ego and Other CAVs** | CAV identity and model | CAV identity and model |
| **Ego and Other CAVs** | Basic Environment Representation | Basic Environment Representation |
| **Ego and Other CAVs** | Full Environment Representation | Full Environment Representation |
| **Ego and Other CAVs** | Messages | Messages |

### Basic Environment Representation

Defined in the Environment Sensing Subsystem.

### CAV Identifier

##### Functionality

A code uniquely identifying a CAV.

##### Requirements

The CAV identification system carries the following information:

1. Country where the CAV has been registered.
2. Registration number in the country.
3. CAV manufacturer identifier.
4. CAV model identifier.

##### Technology

### Commands

##### Functionality

Commands instruct the Motion Actuation Subsystem (MAS) to change its Spatial Attitude SAA at time tA to Spatial Attitude SAB at time tB.

If the Command Issuer signals back to the Obstacle Avoider AIM that there is an anomalous situation, Obstacle Avoider may issue a “discontinue previous Command” and signal back as much as to the Route Planner. This may decide to communicate the situation called “Road State” to the Human-CAV Interaction Subsystem.

##### Requirements

The following data are passed:

1. Duration= TB-TA
2. SAA
3. SAB
4. Environment Data
5. Discontinue Command.

##### Technology

### Feedback

##### Functionality

After receiving a Command from the AMS, the MAS continuously issues Responses back to the AMS to inform about Command execution. The Responses contain the value of a series of SA’sat intermediate Poses and corresponding times until the final Pose and time is reached with SAB at Time TB.

The AMS, based on the Responses received and the potential discovery of changes in the Environment, can decide to discontinue the execution of the earlier command and issue another Command instead.

If the MAS has difficulties execuring the Command it sends back the Time, the SAB, and a code.

##### Requirements

The MAS sends a series of Feedback containing Time and SA. In case of difficulties in exceuting the Command, the following is signaled to the AMS:

1. Estimated inability to reach Goal with the requested SAB at the requested Time TB.
2. Information about Road state
   1. Ice.
   2. Wind.
   3. Water.
   4. Pothole.
3. Wheels or brakes to not respond.

##### Technology

### Road State

##### Functionality

Road State is information about the state of the road the CAV is going through inferred by the AMS from internally available information or received from an external source via a communication channel.

##### Requirements

Road State states are:

1. Road blocked at waypoint wi
2. Traffic jam at waypoint wi
3. The road is slippery at waypoint wi
4. Weather conditions at waypoint wi
   1. Fog at at waypoint x,y,z
5. ...

##### Technology

### Full Environment Representation

##### Functionality

The Autonomous Motion Subsystem produces the Full Environment Representation (FER) by *integrating* information:

1. from the BERs received from other CAVs in range or Roadside Units into the BER received from the Environment Sensing Subsystem.
2. The AMS will reconcile the different values using other considerations, e.g., the distance of the ego CAV from the object vs the distance of the other CAV from the object.
3. Platooning information may be broadcast by the platoon itself or be deduced by the Ego CAV from observation of the behaviour of a group of CAVs over a period of time.

##### Requirements

The FER format shall satisfy the following requirements:

1. The FER shall be an extension of the BER.
2. Each element of the FER shall provide:
   1. The value of major discrepancies between the elements of the Ego CAV’s BER and those of another CAV.
   2. The Identifier, Spatial Attitude, and Shape of the other CAVs.
3. The Objects in the FER may include additional semantics, such as:
   1. Flags (e.g., warning coming from other CAVs and RSUs).
   2. Platooning information indicating which objects belong to a platoon.
   3. Priority (e.g., the vehicle is a police car, an ambulance, a vehicle carrying hazardous material).
4. The FER data structure shall allow AIMs that are inside or connected to the AMS to have fast access to the data structure. For instance, the Obstacle Avoider should have fast access to data in the FER to get information useful for avoiding obstacles to the trajectories of the objects that may hinder the implementation of its decision to reach a given point using a Trajectory.
5. The FER shall allow for deliberative and reactive actions.

##### Technology

### Goal

##### Functionality

A Goal is the Spatial Attitude planned to be reached at the end of a Decision Horizon.

##### Requirements

Goal shall be represented as a Spatial Attitude.

##### Technology

### Offline map

Defined in the Environment Sensing Subsystem.

### Path

##### Functionality

Path is a sequence of Poses 𝑝𝑖 = (xi, yi, zi, αi, βi, γi) in the Offline Map.

##### Requirements

A Path can be represented as a sequence of of Spatial Attitude where the derivatives of Position and Orientation are not included.

##### Technology

### Pose

##### Functionality

Pose is the Position and Orientation of the CAV in the Offline Map p = (𝑥, 𝑦, z, α, β, γ).

The AMS nay issues micro-commands to the MAS in case a Pose cannot be reached from a given Pose in a straight line.

##### Requirements

##### Technology

### Route

##### Functionality

Route is a sequence of Way Points.

##### Requirements

##### Technology

### Traffic rules

##### Functionality

Traffic Rules are digital representations of the traffic rules derived from a set of traffic signals.

##### Requirements

##### Technology

### Traffic Signals

##### Functionality

Traffic Signals are the digital representations of the traffic signals on a road and around it including the semantics of the traffic signs.

##### Requirements

The Traffic Signals Format shall be able to represent:

1. All traffic signalisations required.
2. The specific local version of traffic signalisation.
3. The Position and Orientation of the traffic signals.

##### Technology

### Trajectory

##### Functionality

Trajectory is a sequence of Spatial Attitudes *si* (*s1,s2,…si*) containing the time each Spatial Attitude is expected to be or was reached.

##### Requirements

The Path and the Spatial Attitudes of the CAV at the Poses composing the Trajectory shall be designed so that the time between the beginning and the end of the Trajectory shall not:

1. Require more resources than available in the MAS.
2. Violate the Traffic Rules.
3. Affect passengers’ comfort.

##### Technology

# Motion Actuation Subsystem (MAS)

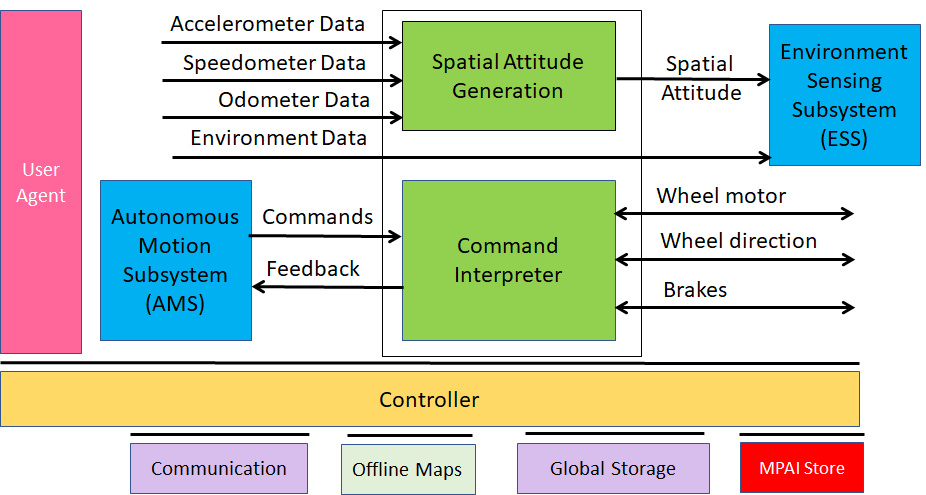
## Use Case description

The Motion Actuation Subsystem (MAS):

1. Transmits information gathered from its sensors and its mechanical subsystems to Environ­ment Sensing Subsystem (ESS).
2. Receives the AMS-MAS Command from Autonomous Motion Subsystem.
3. Translates the AMS-MAS Command into specific commands to its own mechanical subsystems, e.g., wheel direction, wheel motors, and brakes.
4. Receives feedback from its mechanical subsystems.
5. Packages feedback into MAS-AMS Feedback.
6. Send MAS-AMS Feedback to Autonomous Motion Subsystem (AMS).

## Reference architecture

*Figure 8* represents the Motion Actuation Subsystem Reference Model.



*Figure 8 – Motion Actuation Subsystem Reference Model*

## Input and output data

*Table 11* gives the input/output data of Motion Actuation Subsystem.

*Table 11 – I/O data of* *Motion Actuation Subsystem*

|  |  |
| --- | --- |
| **Input** | **Comments** |
| Odometer | Provides distance data. |
| Speedometer | Provides instantaneous velocity. |
| Accelerometer | Provides instantaneous acceleration. |
| Other Environment data | Provide other environment data, e.g., humidity, pressure, temperat­ure. |
| Wheel Motor | Forces road wheels rotation, gives feedback. |
| Wheel Direction | Moves road wheels by an angle, gives feedback. |
| Brakes | Acts on brakes, gives feedback. |
| AMS-MAS Command | High-level motion command from the Autonomous Motion Subsys. |
| **Output** | **Comments** |
| Motion data | Position, velocity, acceleration. |
| Other data | Other environment data. |
| MAS-AMS Feedback | Feedback from Command Interpreter during and after Command ex­ecution. |

## AI Modules

*Table 12* gives the AI Modules of Autonomous Motion Subsystem.

*Table 12 – AI Modules of Motion Actuation Subsystem*

|  |  |
| --- | --- |
| **AIM** | **Function** |
| **Spatial Attitude Generation** | Computes the best estimate of the CAV Spatial Attitude using odometer, speedometer, accelerometer. |
| **Command Interpreter** | Converts AMS-MAS Commands into specific actuation commands to Wheel motors, Wheel Direction and Brakes.  Forwards MAS-AMS Feedback to AMS. |

### Summary of Motion Actuation Subsystem data

*Table 13* gives, for each AIM (1st column), the input data (2nd column) from which AIM (column) and the output data (3rd column).

*Table 13 –Motion Actuation Subsystem data*

|  |  |  |
| --- | --- | --- |
| **CAV/AIM** | **Input** | **Output** |
| **Command Interpreter** | AMS-MAS Command  Wheel Motor Feedback  Wheel Direction Feedback  Brakes Feedback | MAS-AMS Feedback  Wheel Motor Command  Wheel Direction Command  Brakes Command |
| **Spatial Attitude Generation** | Accelerometer  Speedometer  Odometer | Spatial Attitude |

### Accelerometer Data

##### Functionality

The electronic sensor accelerometer measures the acceleration forces acting on the CAV. A CAV accelerometer at rest measures an acceleration g straight upwards of ~9.81 m/s2. If it is in free fall (falling toward the centre of the Earth at ≈ 9.81 m/s2) the acceleration measures zero. Acceleration of a CAV is typically expressed by a vector with values in the three x,y,z coordinates.

##### Requirements

##### Technology

### Brakes Command

##### Functionality

The result of the interpretation of AMS-MAS Command into commands to Brakes. Commands can be expressed as the

##### Requirements

##### Technology

### Brakes Feedback

##### Functionality

The feedback of Brakes to Command Interpreter.

##### Requirements

##### Technology

### Wheel motor Command

##### Functionality

##### Requirements

##### Technology

### Wheel motor Feedback

##### Functionality

##### Requirements

##### Technology

### Wheel Direction Command

##### Functionality

##### Requirements

##### Technology

### Wheel Direction Feedback

##### Functionality

##### Requirements

##### Technology

### AMS-MAS Command

##### Functionality

The Command issued by the AMS

##### Requirements

##### Technology

### MAS-AMS Feedback

##### Functionality

The Feedback of Command Interpreter summarising the Feedbacks.

##### Requirements

##### Technology

### Spatial Attitude

##### Functionality

The MAS generates the Spatial Attitude of the ego CAV using its sensors.

##### Requirements

##### Technology

### Odometer Data

##### Functionality

The odometer converts the number of wheel rotations times the tire circumference (π x tire diameter) from the start up to the point being considered to the distance travelled.

##### Requirements

##### Technology

### Other Environment Data

##### Functionality

The set of Environment data such as temperature, air pressure, humidity etc.

##### Requirements

##### Technology

### Road Wheel Direction Command

##### Functionality

The result of the interpretation of AMS Command to Road Wheel Direction.

##### Requirements

##### Technology

### Road Wheel Direction Feedback

##### Functionality

The feedback of Road Wheel Direction to Command Interpreter.

##### Requirements

##### Technology

### Road Wheel Motor Command

##### Functionality

The result of the interpretation of AMS Command to Road Wheel Motor.

##### Requirements

##### Technology

### Road Wheel Motor Feedback

##### Functionality

The feedback of Road Wheel Motor to Command Interpreter.

##### Requirements

##### Technology

### Speedometer

##### Functionality

An electronic sensor that measures the instantaneous speed of CAV.

##### Requirements

##### Technology

# CAV-to-Everything (V2X)

Communication external to the is handled directly by Subsystems.

## Description

To improve its Environment perception capabilities, CAV exchanges information via radio with other entities, e.g., CAVs in range and other CAV-aware communication devices such as Roadside Units and Traffic Lights in a secure fashion. The current version of this document does not consider the secure CA-to-CAV communication.

A CAV may broadcast to CAVs in range information it has become aware of by using the V2X communication interface. For instance, while executing a Command, the MAS of CAV may become aware of ice on the road. The CAV may decide to broadcast that information to CAVs in range.

Multicast Communication is used when CAV broadcasts its identity and in case of heavy data types, e.g., Basic Environment Representation (BER). Unicast mode may be used in other cases.

A Communication Device outside of the MPAI-AIF Trusted Zone of the Autonomous Motion Subsystem (AMS) in in charge of communication. The Device communicates with any AIF of the CAV which has communication needs or from which the Communication Device has received data.

The flow of operations of the Communication Device when handling communication with other CAVs or with devices having CAV functionality (e.g., a traffic light or a roadside unit) is:

1. Receives identities broadcasted by CAVs in range.
2. Establishes unicast sessions with CAVs in range. Issues to be considered are:
   1. Can CAV’s Communication Device handle a large number of sessions?
   2. Can the AMS process data from a large number of different CAVs?
   3. Which application protocol is used by the Communication Device?
   4. Does the delay inherent in a unicast protocol have an impact?
3. Creates a list of CAVs in range with which it has established a session.
4. Sends the list with Basic Environment Representations (BER) received via broadcast to the Autonomous Motion Subsystem (AMS).
5. ESS sends its own BER to the Communication Device.
6. Communication Device broadcasts BER in encrypted form with a key that is only known to CAVs in range that have an open unicast session with the Communication Device.

Note: no decision has been made on whether CAV should send/receive Full Environment Representations (FER).

The Communication Device is also made aware of the nature of the CAV and CAV-like device:

1. Traffic light.
2. Fire Truck.
3. Police.
4. Ambulance.
5. Flock Leader.

CAVs should communicate using a protocol that assigns a slice of the available transmission rate to each CAV based on the number of CAVs.

## Input and output data

### CAVs within range

*Table 14* gives the data types CAV broadcasts to CAVs in range via its Communication Device.

*Table 14 – I/O data of* *CAV’s Communication Device*

|  |  |  |
| --- | --- | --- |
| **Input Data** | **From** | **Comments** |
| Basic Environment Representation | Other CAVs | A digital representation of the Environment created by the ESS. |
| CAV Identity | Other CAVs | In principle, this should be the digital equivalent of today’s plate number including Manufacturer and Model information. The need to share this information is TBD. |
| CAV Intention | Other CAVs | The Path and other motion data relevant to other CAVs |
| Full Environment Representation | Other CAVs | A digital representation of the Environment created by fusing all available Basic Environment Representations. The need to share this information is TBD. |
| Information Messages | Other CAVs | Typical messages CAV can broadcast. Potentially important messages for CAVs are given by [36, 37]   1. CAV is an ambulance. 2. CAV carries an authority. 3. CAV carries a passenger with critical health problem. 4. CAV has a mechanical problem of an identified level. 5. Works and traffic jams ahead 6. Environment must be evacuated 7. .... |
| **Output Data** | **To** | **Comments** |
| Basic Environment Representation | Other CAVs | Same as input for all other input data. |
| Full Environment Representation | Other CAVs | A digital representation of the Environment obtained from the fusion of all available Basic Environment Representations. The need to share this information is TBD. |

### CAV-aware equipment

Examples of such equipment are traffic lights, roadside units, vehicles with CAV communication capabilities. The following data may be exchanged:

1. Identity and coordinates (exact coordinate reference).

Static Full Environment Representation regularly updated via download (may be part of the Offline Map).

Current objects in Environment.

State (Green-Yellow-Red) of traffic light and time to change state.

Lane markings.

Speed limits.

Pedestrian crosswalks

General information on the Environment (e.g., one way street etc.)

Etc.

Such equipment can:

1. Act as any other CAV in range.
2. Have the authority to organise motion of CAVs in range.

### Other non-CAV vehicles

Other vehicles can be scooters, motorcycles, bicycles, etc. possibly transmitting their position as derived from GNSS. No response capability is expected. Vehicles may also have the capability to transmit additional information, e.g., identity, model, speed.

### Pedestrians

Their smartphones can transmit their coordinates as available from GNSS. No response capability is expected.

# Data privacy

CAV can generate or acquire data for which privacy is an important characteristic. Here are some of the functions potentially affected by data privacy or that are liable to become accessible to authorities, e.g., police and judiciary.

## Human-CAV Interaction (HCI)

By having interactions with humans, HCI becomes aware of potentially sensitive information, e.g.:

1. Result of monitoring the passenger cabin.
2. Minute requests from humans, e.g., go to a way point, display Full Environment Representation, turn off air conditioning, etc.
3. Dialogue with human

## Environment Sensing Subsystem (ESS)

ESS collects large among of environment data for the purpose of creating instantaneous Basic Environment Representations, e.g.:

1. GNSS gives the position of the CAV and of whatever is perceived by the CAV that is approximate, but sufficiently precise for the intended uses.
2. Radar, Lidar, Ultrasound give variously defined information about what is in the environment surrounding the CAV.
3. Cameras give a 360° panoramic view of the environment where all objects, save those occluded, are visible.
4. External microphones give a complete representation of the external sound field.

A user could create a permanent and certified recording of important data acquired by ESS.

The environment recorder could compress and record all data acquired for a limited amount of time. Some data could be recorded for a longer time.

## CAV to Everything (V2X)

V2X acquires the identity of the CAVs in range and communicates appropriate subsets of the Basic and Full Environment s Representations.

## Autonomous motion subsystem (AMS)

AMS knows the exact waypoints the CAV has passed through and all the commands given to the Motion Actuation Subsystem.

By integrating the Basic Environment Representations of all CAVs in range and its own, CAV can create a detailed and extended map of the environment.

Recording the decisions made by the Decision Recorder creates highly critical data.

## Motion Actuation Subsystem (MAS)

MAS acquires position information through its Spatial Attitude Generation AIM.

1. – General MPAI Terminology

The Terms used in this standard whose first letter is capital and are not already included in *Table 1* are defined in *Table 15.*

*Table 15 – 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 AIM-specific inputs and producing AIM-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 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” and “Conformance Testing” and, where applicable, “Performance Testing”.

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 especially established MPAI Development Committees who operate on the basis of 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 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](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**

With reference to *Figure 1*, 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 be an Implementation of 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 | * 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 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**

The following *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 [10]:

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 MPAI Store.
6. Implementers submit Implementations to the MPAI Store
7. MPAI Store verifies security and Tests Confor­mance of Implementation.
8. Users download Implementations and report their experience to MPAI.

Diagram

Description automatically generated

*Figure 9 – The MPAI ecosystem operation*

## Informative References

This document references the following informative documents:

1. Technical Specification: The Governance of the MPAI Ecosystem (MPAI-GME) V1.1; https://mpai.community/standards/mpai-gme/.
2. Global Autonomous Cars (Semi & Fully) Market Report 2021, https://www.prnewswire.com/news-releases/global-autonomous-cars-semi--fully-market-report-2021-market-is-expected-to-reach-1-383-89-billion-in-2025---forecast-to-2030--301292821.html
3. Serban, Alexandru Constantin, Erik Poll, and Joost Visser. “A Standard Driven Software Architecture for Fully Autonomous Vehicles.” 2018 IEEE International Conference on Software Architecture Companion (ICSA-C). IEEE, 2018.
4. P. Avgeriou, "Describing, Instantiating and Evaluating a Reference Architecture: A Case Study," Enterprise Architecture Journal, June 2003.
5. S. Behere and M. Törngren, "A functional reference architecture for autonomous driving," Information and Software Technology 73, 2016, 136-150.
6. K. Jo, J. Kim, D. Kim, C. Jang, and M. Sunwoo, "Development of Autonomous Car - Part II: A Case Study on the Implementation of an Autonomous Driving System Based on Distributed Architecture," IEEE Transactions on Industrial Electronics, 62(8), 2015, 5119-5132.
7. K. Jo, M. Lee, D. Kim, J. Kim, C. Jang, E. Kim, and M. Sunwoo, "Overall Reviews of Autonomous Vehicle A1-System
8. Architecture and Algorithms," The International Federation of Automatic Control, 46(10), 2013, 114-119.
9. C. Maple, M. Bradbury, A. T. Le, and K. Ghirardello, "A Connected and Autonomous Vehicle Reference Architecture for Attack Surface Analysis," Applied Sciences, 9(23), 2019. doi:10.3390/app9235101
10. U. S. Department of Transportation, "Architecture Reference for Cooperative and Intelligent Transportation,", 2019, Retrieved March 03, 2020 from https://local.iteris.com/arc-it/html/architecture/architecture.html
11. SAE International Releases Updated Visual Chart for Its “Levels of Driving Automation” Standard for Self-Driving Vehicles, https://www.sae.org/news/press-room/2018/12/sae-international-releases-updated-visual-chart-for-its-%E2%80%9Clevels-of-driving-automation%E2%80%9D-standard-for-self-driving-vehicles
12. ISO 8855: "Road vehicles -- Vehicle dynamics and road-holding ability -- Vocabulary"
13. Rodolfo W. L. Coutinho and Azzedine Boukerche, Guidelines for the Design of Vehicular Cloud Infrastructures for Connected Autonomous Vehicles, IEEE Wireless Communications - August 2019
14. Claudine Badue, Rânik Guidolini, Raphael Vivacqua Carneiro, Pedro Azevedo, Vinicius B. Cardoso, Avelino Forechi, Luan Jesus, Rodrigo Berriel, Thiago M. Paixão, Filipe Mutz, Lucas de Paula Veronese, Thiago Oliveira-Santos, Alberto F. De Souza; Self-driving cars: A survey; Expert Systems With Applications 165 (2021) 113816
15. D. Cireşan, U. Meier, J. Masci, and J. Schmidhuber, “Multi-column deep neural network for traffic sign classification,” Neural Netw., vol.32, pp.333–338, Aug. 2012
16. ETSI TR 103 562 V2.1.1 (2019-12), Analysis of the Collective Perception Service (CPS); Release 2.
17. Gokulnath Thandavarayan, Miguel Sepulcre, and Javier Gozalvez; Generation of Cooperative Perception Messages for Connected and Automated Vehicles; IEEE Transactions on Vehicular Technology, Vol. 69, No. 12, December 2020
18. CAR 2 CAR Communication Consortium, https://www.car-2-car.org/
19. Usman Ali Khan and Sang Sun Lee; Distance-Based Resource Allocation for Vehicle-to-Pedestrian Safety Communication; https://www.mdpi.com/2079-9292/9/10/1640/pdf
20. Pranav Kumar Singhab, Sunit Kumar Nandiac, Sukumar Nandi; A tutorial survey on vehicular communication state of the art, and future research directions; Vehicular Communications Volume 18, August 2019, 100164
21. https://phantom.ai/assets/uploads/PAI%20Renesas%20Partnership%20Announcement%20(1).pdf
22. Charles R., Qi Li, Yi Hao Su, Leonidas J. Guibas; PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space; <https://arxiv.org/pdf/1706.02413.pdf>
23. Bernard, Olivier; Bradway, David; Hansen, Hendrik H.G.; Kruizinga, Pieter; Nair, Arun; Perdios, Dimitris; Ricci, Stefano; Rindal, Ole Marius Hoel; Rodriguez-Molares, Alfonso; Stuart, Matthias Bo; The Ultrasound File Format (UFF) - First draft, Proceedings of 2018 IEEE International Ultrasonics Symposium
24. LAS (LASer) File Format, Version 1.4 ,  
     https://www.loc.gov/preservation/digital/formats/fdd/fdd000418.shtml
25. https://www.lesliesikos.com/pcap/
26. https://pointclouds.org/documentation/tutorials/hdl\_grabber.html
27. https://www.mathworks.com/help/vision/ref/velodynefilereader.html
28. M. Heistermann, S. Jacobi and T. Pfaff,Technical Note: An open source library for processing weather radar data (wradlib), https://hess.copernicus.org/articles/17/863/2013/hess-17-863-2013.pdf
29. https://pro.arcgis.com/en/pro-app/latest/help/data/data-interoperability/supported-formats-with-the-data-interoperability-extension.htm
30. CDOT and Panasonic Take First Steps to Turn I-70 into Connected Roadway, <https://www.codot.gov/news/2018/july/cdot-and-panasonic-take-first-steps-to-turn-i-70-into-connected-roadway>
31. Navigation Data Standards, https://nds-association.org/
32. Matthias Schreier; Environment Representations for Automated On-Road Vehicles; https://www.researchgate.net/publication/323105152\_Environment\_Representations\_for\_Automated\_On-Road\_Vehicles
33. Toh Chai Keong, Juan Carlos Cano, Carlos-Javier Fernandez-Laguia, Pietro Manzoni; Wireless Digital Traffic Signs of the Future; IET Networks, September 2018
34. Vladimir Hahanov, Wajeb Gharibi, Eugenia Litvinova, Svitlana Chumachenko, Arthur Ziarmand, Irina Englesi, Igor Gritsuk, Vladimir Volkov, Anastasiia Khakhanova; Cloud-Driven Traffic Monitoring and Control Based on Smart Virtual Infrastructure; 2017-03-28
35. SharedStreets; <https://sharedstreets.io/>
36. Bojarski, Mariusz, et al. “End to end learning for self-driving cars.” arXiv preprint arXiv:1604.07316 (2016).
37. James S. Albus; The NIST Real-time Control System (RCS) An Approach to Intelligent Systems Research; <https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=820528>
38. Integrated Public Alert & Warning System; <https://www.fema.gov/gu/emergency-managers/practitioners/integrated-public-alert-warning-system/public>
39. OASIS Standard; Common Alerting Protocol Version 1.2; <https://docs.oasis-open.org/emergency/cap/v1.2/CAP-v1.2-os.pdf>
40. Dorais G., Kortenkamp D. (2001) Designing Human-Centered Autonomous Agents. Lecture Notes in Computer Science, vol 2112. Springer, Berlin, Heidelberg. <https://doi.org/10.1007/3-540-45408-X_32>
41. <https://youtu.be/nCaomvjPIok>
42. Stefano Mariani and Franco Zambonelli; Degrees of Autonomy in Coordinating Collectives of Self-Driving Vehicles; ISoLA 2021
43. Automatic Radar Performance Metrics and Testing Methods for Automated Driver Assistance Systems (ADAS) and Automated Driving Systems Applications (ADS); IEEE P3116; <https://sagroups.ieee.org/3116/>
44. Paulitsch, H., Fuchsberger, J., and Kock, K.: FM94-BUFR Encoding and Decoding Software, User Guidelines Version 1.6 For BUFR Software Version 3, http://www.knmi.nl/opera/bufr/doc/bufr\_sw\_desc.pdf, 2010.
45. Michelson, D. B., Lewandowski, R., Szewczykowski, M. and Beekhuis, H. (2011). EUMETNET OPERA weather radar information model for implementation with the HDF5 file format, version 2.1, Opera working document WD\_2008\_03, EUMETNET OPERA.
46. MPAI; MPAI-MMC V2 Use Cases and Functional Requirements; <https://mpai.community/standards/mpai-mmc/about-mpai-mmc/#V2>.
47. Intelligent Transport Systems (ITS); Facilities Layer function; Part 2: Position and Time management (PoTi); Release 2; ETSI EN 302 890-2.
48. Letizia Marchegiani, Xenofon Fafoutis; How Well Can Driverless Vehicles Hear? An Introduction to Auditory Perception for Autonomous and Smart Vehicles; IEEE Intelligent Transportation Systems Magazine; May/June 2022; <https://ieeexplore.ieee.org/document/9352026?mkt_tok=NzU2LUdQSC04OTkAAAGF_oHR8sRXCaOYj8X4r3C6PmWYuuhBVayVAyD0AGTR_-JhPa-VZS0LKHhpACcNGxFd7qmkZ83XnBW_DwGR8SQeXnmJOoIpkEqrxsjN1UzOSI3WMg>

1. At the time of publication of this Technical Report, the MPAI Store was assigned as the IIDRA. [↑](#footnote-ref-1)
2. A time-of-flight device measures the distance and speed of a sensed object based on the time it takes for a signal to hit an object and be reflected. [↑](#footnote-ref-2)
3. Traffic rules are derived from the semantics of the objects, e.g., speed limit 50 km/h, not from the reasoning made on the objects which this document assumes it is within the scope of the Autonomous Motion Subsystem. [↑](#footnote-ref-3)