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

|  |  |
| --- | --- |
| **Public Document** | |
| **N410** | 2021/10/27 |
| **Source** | Requirements (CAV) |
| **Title** | MPAI-CAV Use Cases and Functional Requirements WD0.5 |
| **Target** | MPAI-13 |

[1 Introduction 3](#_Toc86238978)

[2 Scope of the MPAI-CAV Use Cases 4](#_Toc86238979)

[3 Terms and definitions 5](#_Toc86238980)

[4 References 7](#_Toc86238981)

[4.1 Normative References 7](#_Toc86238982)

[4.2 Informative References 7](#_Toc86238983)

[5 Use Cases 8](#_Toc86238984)

[5.1 Human-CAV Interaction (HCI) 10](#_Toc86238985)

[5.1.1 Use Case description 10](#_Toc86238986)

[5.1.2 Reference architecture 11](#_Toc86238987)

[5.1.3 Input and output data 12](#_Toc86238988)

[5.1.4 AI Modules 12](#_Toc86238989)

[5.2 Environment Sensing Subsystem (ESS) 13](#_Toc86238990)

[5.2.1 Use Case description 13](#_Toc86238991)

[5.2.2 Reference architecture 13](#_Toc86238992)

[5.2.3 Input and output data 14](#_Toc86238993)

[5.2.4 AI Modules 14](#_Toc86238994)

[5.3 CAV-to-Everything (V2X) 15](#_Toc86238995)

[5.3.1 Use Case description 15](#_Toc86238996)

[5.3.2 Reference architecture 15](#_Toc86238997)

[5.3.3 Input and output data 16](#_Toc86238998)

[5.3.4 AI Modules 17](#_Toc86238999)

[5.4 Autonomous Motion Subsystem (AMS) 17](#_Toc86239000)

[5.4.1 Use Case description 17](#_Toc86239001)

[5.4.2 Reference architecture 18](#_Toc86239002)

[5.4.3 Input and output data 18](#_Toc86239003)

[5.4.4 AI Modules 19](#_Toc86239004)

[5.5 Motion Actuation Subsystem (MAS) 19](#_Toc86239005)

[5.5.1 Use Case description 19](#_Toc86239006)

[5.5.2 Reference architecture 20](#_Toc86239007)

[5.5.3 Input and output data 20](#_Toc86239008)

[5.5.4 AI Modules 20](#_Toc86239009)

[6 Functional Requirements 21](#_Toc86239010)

[6.1 Human-CAV Interaction 21](#_Toc86239011)

[6.1.1 I/O Data summary 21](#_Toc86239012)

[6.1.2 Audio 22](#_Toc86239013)

[6.1.3 Verbal Interaction 22](#_Toc86239014)

[6.1.4 Concept Expression (Face) 23](#_Toc86239015)

[6.1.5 Concept Expression (Speech) 23](#_Toc86239016)

[6.1.6 Emotion 23](#_Toc86239017)

[6.1.7 Face identity 23](#_Toc86239018)

[6.1.8 Face Objects 23](#_Toc86239019)

[6.1.9 Full World Representation 24](#_Toc86239020)

[6.1.10 Full World Representation Viewer commands 24](#_Toc86239021)

[6.1.11 Intention 24](#_Toc86239022)

[6.1.12 Meaning 24](#_Toc86239023)

[6.1.13 Object Identifier 24](#_Toc86239024)

[6.1.14 Speaker Identity 24](#_Toc86239025)

[6.1.15 Text 25](#_Toc86239026)

[6.1.16 Video 25](#_Toc86239027)

[6.2 Environment Sensing Subsystem 25](#_Toc86239028)

[6.2.1 I/O Data summary 25](#_Toc86239029)

[6.2.2 Microphone Array Audio 26](#_Toc86239030)

[6.2.3 Audio Objects 26](#_Toc86239031)

[6.2.4 Basic World Representation 26](#_Toc86239032)

[6.2.5 GNSS Coordinates 27](#_Toc86239033)

[6.2.6 GNSS Data 27](#_Toc86239034)

[6.2.7 Lidar Data 27](#_Toc86239035)

[6.2.8 Moving Object Tracker Data 28](#_Toc86239036)

[6.2.9 Offline Maps 28](#_Toc86239037)

[6.2.10 Radar Data 28](#_Toc86239038)

[6.2.11 State 29](#_Toc86239039)

[6.2.12 Traffic Signalisation 29](#_Toc86239040)

[6.2.13 Ultrasound Data 29](#_Toc86239041)

[6.2.14 Video Camera data 30](#_Toc86239042)

[6.2.15 Visual Objects and Scene (Camera) 30](#_Toc86239043)

[6.2.16 Visual Objects and Scene (Lidar) 30](#_Toc86239044)

[6.2.17 Visual Objects and Scene (Radar) 30](#_Toc86239045)

[6.2.18 Ultrasound Objects and Scene (Ultrasound) 30](#_Toc86239046)

[6.3 CAV to Everything 30](#_Toc86239047)

[6.3.1 Summary of CAV to Everything data 30](#_Toc86239048)

[6.3.2 Basic World Representation 31](#_Toc86239049)

[6.3.3 CAV Identifier 31](#_Toc86239050)

[6.3.4 Events 31](#_Toc86239051)

[6.3.5 Full World Representation 31](#_Toc86239052)

[6.3.6 Path 32](#_Toc86239053)

[6.3.7 State 32](#_Toc86239054)

[6.3.8 Trajectory 32](#_Toc86239055)

[6.4 Autonomous Motion Subsystem 32](#_Toc86239056)

[6.4.1 Summary of Autonomous Motion Subsystem data 32](#_Toc86239057)

[6.4.2 Basic World Representation 32](#_Toc86239058)

[6.4.3 Command/Response 33](#_Toc86239059)

[6.4.4 Full World Representation 33](#_Toc86239060)

[6.4.5 Goal 33](#_Toc86239061)

[6.4.6 Offline map 33](#_Toc86239062)

[6.4.7 Path 33](#_Toc86239063)

[6.4.8 Pose 33](#_Toc86239064)

[6.4.9 Route 33](#_Toc86239065)

[6.4.10 State 34](#_Toc86239066)

[6.4.11 Traffic rules 34](#_Toc86239067)

[6.4.12 Traffic Signals 34](#_Toc86239068)

[6.4.13 Trajectory 34](#_Toc86239069)

[6.4.14 Velocity 34](#_Toc86239070)

[6.5 Motion Actuation Subsystem 34](#_Toc86239071)

[6.5.1 Summary of Motion Actuation Subsystem data 34](#_Toc86239072)

[6.5.2 Accelerometer data 35](#_Toc86239073)

[6.5.3 Brakes Command 35](#_Toc86239074)

[6.5.4 Brakes Feedback 35](#_Toc86239075)

[6.5.5 Command from AMS 35](#_Toc86239076)

[6.5.6 Feedback to AMS 35](#_Toc86239077)

[6.5.7 Motion Data 35](#_Toc86239078)

[6.5.8 Odometer Data 35](#_Toc86239079)

[6.5.9 Other Environment Data 36](#_Toc86239080)

[6.5.10 Road Wheel Direction Command 36](#_Toc86239081)

[6.5.11 Road Wheel Direction Feedback 36](#_Toc86239082)

[6.5.12 Road Wheel Motor Command 36](#_Toc86239083)

[6.5.13 Road Wheel Motor Feedback 36](#_Toc86239084)

[6.5.14 Speedometer 36](#_Toc86239085)

[7 Data privacy 36](#_Toc86239086)

[7.1 Human-CAV Interaction (HCI) 36](#_Toc86239087)

[7.2 Environment Sensing Subsystem (ESS) 37](#_Toc86239088)

[7.3 CAV to Everything (V2X) 37](#_Toc86239089)

[7.4 Autonomous motion subsystem (AMS) 37](#_Toc86239090)

[7.5 Motion Actuation Subsystem (MAS) 37](#_Toc86239091)

[Annex 1 – General MPAI Terminology (Normative) 38](#_Toc86239092)

[Annex 2 - Notices and Disclaimers Concerning MPAI Standards (Informative) 41](#_Toc86239093)

[Annex 3 – The Governance of the MPAI Ecosystem (Informative) 43](#_Toc86239094)

[Annex 4 – Datasets for CAV research 45](#_Toc86239095)

[Annex 5 – ETSI Technical Report 46](#_Toc86239096)

[Annex 6 – Some CAV Communication Technologies 47](#_Toc86239097)

# Introduction

Moving Picture, Audio and Data Coding by Artificial Intelligence (MPAI) is an [international Standards Developing Organisation](http://mpai.community/) with the mission to develop *AI-enabled data coding standards*. Research has shown that data coding with AI-based technologies is generally *more efficient* than with existing technologies. Compression and feature-based description are notable examples of coding.

In the following, Terms beginning with a capital letter are defined in *Table 3* if they are specific to MPAI-MCS Standard and to *Table 17* if they are common to all MPAI Standards.

MPAI Application Standards enable the development of AI-based products, applic­ations and services. The MPAI AI Framework (AIF) Standard (MPA-AIF) [2] provides the foundation on which the technologies defined by MPAI Application Standards operate.

*Figure 1* depicts the MPAI-AIF Reference Model. This Introduction only describes the basic proc­essing elements called AI Modules (AIM) which make up an AI Workflow (AIW) executed in an AI Framework (AIF).

Diagram

Description automatically generated

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

MPAI Application Standards normatively specify:

1. *For the AIMs*: the Function, and the Semantics and Formats of the input and output data but not the internal architecture, which may be based on AI or data processing, and implemented in software, hardware or hybrid software and hardware technologies.
2. *For the AIWs*: the Function, the input and output data Semantics and Formats, and the Topology of the Connections between and among the AIMs.

MPAI defines Interoperability as the ability to replace an AIW or an AIM Implementation with a functionally equivalent AIW or AIM Implementation. An AIW *executed in an AIF* may have one of the following MPAI-defined Interoperability Levels:

1. *Interoperability Level 1*, if the AIW is proprietary and composed of AIMs with proprietary functions using any proprietary or standard data Format.
2. *Interoperability Level 2*, if the AIW is composed of AIMs having all their Functions, Formats and Connections specified by an MPAI Application Standard.
3. *Interoperability Level 3*, if the AIW has Interoperability Level 2, and the AIW and its AIMs are certified by an MPAI-appointed Assessor to hold the attributes of Rel­iability, Robustness, Replicability and Fairness – collectively called Performance.

MPAI is the basic element of the MPAI Ecosystem [1] offering 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.

# Scope of the MPAI-CAV Use Cases

Standardisation of Connected Autonomous Vehicles CAV components will be required because of: 1) the different nature of the interacting technologies making up a CAV, 2) the sheer size of the future CAV market and 3) the need for users and regulators alike to be assured of CAV safety, reliability and explainability.

At this point in time, a traditional approach to standardisation might consider CAV standards premature and some affected industries may not even be ready yet to consider CAV standards. CAVs, however, at best belong to an industry still being formed, expected to target the production of economic affordable units in the hundreds of millions, with components to be produced by disparate sources. A competitive market of standard components can reduce costs and make CAV confirm their promise to have a major positive impact on the environment and society.

A CAV Reference Model (RM) identifying components and their interfaces can accelerate the definition of standard components. Progression from research to standardation can unfold as a series of proposals from research suggesting components and interfaces to standardisation, and standardisation either requesting more results, or refining the results, or adopting the proposals. Eventually, industry will receive a set of specifications for standard component functions and interfaces to be implemented as best available technology allows. Implementation in products will rely, as a minimum, on the know how of those who have driven the development of the specific­ations.

*Connected Autonomous Vehicles* (MPAI-CAV) is an MPAI standard project, comprising sev­eral identified candidate Use Cases.

The MPAI-CAV Reference Model is subdivided in 5 subsystems:

1. *Human-CAV interaction* (HCI) handling human-CAV interactions.
2. *Environment Sensing Subsystem* (ESS) acquiring information from the physical environment via a variety of sensor.
3. *CAV to Everything* (V2X) receiving information from external sources.
4. *Autonomous Motion Subsystem* (AMS) giving commands to drive the CAV to the intended destination.
5. *Motion Actuation Subsystem* (MAS) providing environment information¸ and receiving/a­ctu­ating motion commands in the physical world.

Each of the 5 subsystems is an instantiation of the MPAI-defined AI Framework (AIF).

The Reference Model identifies and describes the requirements of the data types received or generated by the AIMs in each subsystem. It allows researchers to select data, define testing setups, propose update of interfaces, conduct contests, consider the influence of external components, and subdivide workload in a way that allows unambiguous comparison of results.

Unlike previously published papers (e.g., [10]), this document has the following features:

1. A holistic approach that includes all IT components of a CAV;
2. AIF-AIW-AIM as the unifying model to determine the functionality and the data of all CAV components;
3. Modules with functions and data being or already specified in other MPAI standards;
4. Focuses on the data formats between AIMs rather than focus on the modules whose internals are not part of a standard but left to proprietary implementations.
5. A process where research is seamlessly integrated with a subsequent standardisation process.

The purpose of this document is:

1. To collect and describe the identified use cases.
2. To identify the functions, and the input and output data of the AIWs that implement the Use Cases.
3. To identify the connections of the AIMs making up the AIWs.
4. To identify the functions, and the input and output data of the AIMs required to realise the AIWs.

Chapter 6 Provides the functional requirements that the data formats identified in points 3. and 4. above should satisfy and the connections identified in point 3 above.

# Terms and definitions

*Table 1* defines the terms used in this document. Terms are organised by the CAV Subsystems identified in *Figure 3*. The general MPAI Terms are defined in *Table 17*.

*Table 1 – Definition of Terms used in this document organised by Subsystems*

|  |  |  |
| --- | --- | --- |
| Legend | AMS | Autonomous Motion Subsystem |
|  | CAV | Connected Autonomous Vehicle |
|  | ESS | Environment Sensing Subsystem |
|  | HCI | Human-CAV Interaction |
|  | MAS | Motion Actuation Subsystem |
|  | V2X | CAV to Everything |

|  |  |  |
| --- | --- | --- |
| **SubS** | **Term** | **Definition** |
| AMS | Command | High-level instructions whose execution allows a CAV to reach a Goal. |
| AMS | Decision Horizon | The estimated time between the current State and the Goal. |
| AMS | Full World Representation | A digital representation of the Environment obtained from the fusion of all available Basic World Representations. |
| AMS | Goal | The planned State at the end of the Decision Horizon. |
| AMS | Full World Representation | A description of Environment using the CAV’s and other CAVs’ Basic World Representation. |
| AMS | Path | A sequence of Poses 𝑝𝑖 = (𝑥𝑖,𝑦𝑖,zi,𝜃𝑖) in the Offline Map. |
| AMS | Pose | Coordinates and orientation of the CAV in the Offline Map *p* = (𝑥,𝑦,z,𝜃) |
| AMS | Route | A sequence of Way Points |
| AMS | State | CAV’s Pose, Velocity and Acceleration at a given time. |
| AMS | Traffic Rules | The digital representation of the traffic rules applying to a Pose. |
| AMS | Way Point | A point 𝑤𝑖 given as a coordinate pair (𝑥𝑖, 𝑦𝑖), in an Offline Map |
| CAV | Connected Autonomous Vehicle | A vehicle capable to autonomously reach an assigned target by understanding human utterances, planning a route, sensing and interpreting the environment, exchanging information with other CAVs and acting on the CAV’s motion subsystem. |
| CAV | Health | The condition, e.g., mechanical, of a Subsystem or an AIM. |
| CAV | Reference Model | The collection of the following resources:   1. AIW input and output data. 2. Connections of the AIMs. 3. AIMs’ input and output data. 4. AIMs’ connections. |
| CAV | Subsystem | One of the 5 components making up the CAV. |
| ESS | Basic World Representation | A digital representation of the Environment created with infor­mation available from the CAV’s ESS and an Offline Map. |
| ESS | Environment | The portion of the world of interest to the CAV. |
| ESS | Global Navigation Satellite System | (GNSS) includes 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., accelerometer, speedometer, gyroscope, odometer etc. |
| ESS | Offline Map | An offline-created map of a location and associated metadata. |
| HCI | Command | High-level instructions whose execution allows a CAV to reach a Goal. |
| HCI | Feedback | Responses or sentences autonomously generated by the CAV. |
| MAS |  |  |
| V2X |  |  |

# References

## Normative References

This document references the following normative documents:

1. Technical Specification: The Governance of the MPAI Ecosystem V1.
2. Technical Specification: AI Framework (MPAI-AIF), MPAI document N359.
3. Technical Specification: Technical Specification: Multimodal Conversation (MPAI-MMC) V1; published at https://mpai.community/standards/resources/1.
4. Draft Technical Specification: Context-based Audio Enhancement (MPAI-CAE) VI, to be published at https://mpai.community/standards/resources/.
5. Universal Coded Character Set (UCS): ISO/IEC 10646; December 2020
6. ISO/IEC 14496-10; Information technology – Coding of audio-visual objects – Part 10: Advanced Video Coding.
7. ISO/IEC 23008-2; Information technology – High efficiency coding and media delivery in heterogeneous environments – Part 2: High Efficiency Video Coding.
8. ISO/IEC 23094-1; Information technology – General video coding – Part 1: Essential Video Coding.

## Informative References

This document references the following informative documents:

1. 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
2. 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.
3. ISO 8855: "Road vehicles -- Vehicle dynamics and road-holding ability -- Vocabulary"
4. Rodolfo W. L. Coutinho and Azzedine Boukerche, Guidelines for the Design of Vehicular Cloud Infrastructures for Connected Autonomous Vehicles, IEEE Wireless Communications - August 2019
5. 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
6. 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
7. ETSI TR 103 562 V2.1.1 (2019-12), Analysis of the Collective Perception Service (CPS); Release 2.
8. 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
9. CAR 2 CAR Communication Consortium, https://www.car-2-car.org/
10. 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
11. 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
12. https://phantom.ai/assets/uploads/PAI%20Renesas%20Partnership%20Announcement%20(1).pdf
13. 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>
14. 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
15. LAS (LASer) File Format, Version 1.4 , https://www.loc.gov/preservation/digital/formats/fdd/fdd000418.shtml
16. https://www.lesliesikos.com/pcap/
17. https://pointclouds.org/documentation/tutorials/hdl\_grabber.html
18. https://www.mathworks.com/help/vision/ref/velodynefilereader.html
19. 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
20. https://pro.arcgis.com/en/pro-app/latest/help/data/data-interoperability/supported-formats-with-the-data-interoperability-extension.htm
21. 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>
22. Navigation Data Standards, https://nds-association.org/
23. Matthias Schreier; Environment Representations for Automated On-Road Vehicles; https://www.researchgate.net/publication/323105152\_Environment\_Representations\_for\_Automated\_On-Road\_Vehicles
24. Toh Chai Keong, Juan Carlos Cano, Carlos-Javier Fernandez-Laguia, Pietro Manzoni; Wireless Digital Traffic Signs of the Future; IET Networks, September 2018
25. 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
26. SharedStreets; <https://sharedstreets.io/>
27. Bojarski, Mariusz, et al. “End to end learning for self-driving cars.” arXiv preprint arXiv:1604.07316 (2016).
28. 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>
29. Integrated Public Alert & Warning System; <https://www.fema.gov/gu/emergency-managers/practitioners/integrated-public-alert-warning-system/public>
30. OASIS Standard; Common Alerting Protocol Version 1.2; <https://docs.oasis-open.org/emergency/cap/v1.2/CAP-v1.2-os.pdf>
31. 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>
32. <https://youtu.be/nCaomvjPIok>

# Use Cases

MPAI-CAV seeks to standardise all components that enable the implementation of a Connected Autonomous Vehicle (CAV), i.e., a mechanical system capable of executing the command to move its body autonomously – save for the exceptional intervention of a human – based on the analysis of the data produced by a range of sensors exploring the environment and the information trans­mitted by other sources in range, e.g., CAVs and roadside units (RSU).

*Figure 2* depicts the context where a CAV operates.

A picture containing icon

Description automatically generated

*Figure 2 – The environment where a CAV operates*

MPAI-CAV includes 5 Use Cases that correspond to the 5 main subsystems of a Connected Autonomous Vehicle:

*Human-CAV interaction (HCI)* recognises the human CAV rights holder, responds to humans’ commands and queries, provides extended environment representation (Full World Repres­entation) for humans to enjoy, senses human activities during the travel and may activate other subsystems as required by humans or as deemed necessary by the identified conditions.

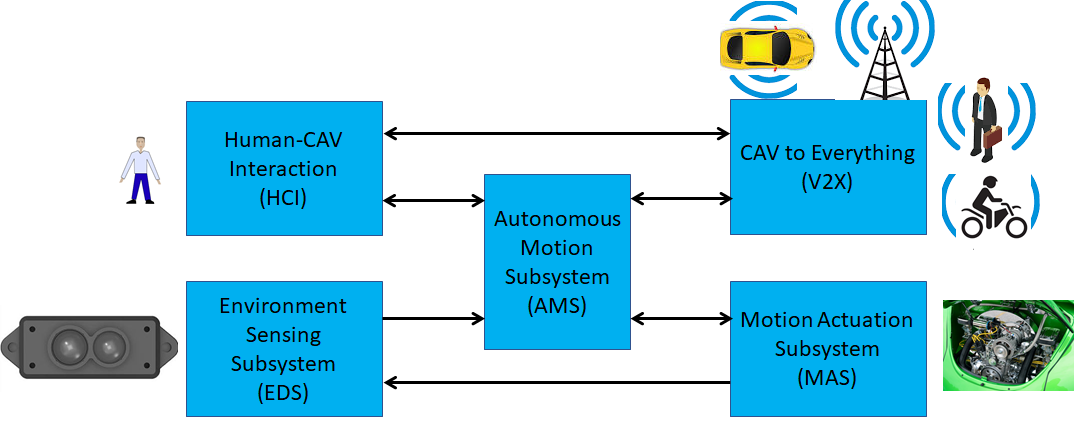
*Environment Sensing Subsystem (EDS)* acquires information from the physical environment via a variety of sensors, develops the best representation of the environment (Basic World Representation) and transfers it to Autonomous Motion Subsystem.

*CAV to Everything Subsystem (V2X)* sends/receives information to/from external sources, including other communicating entities in range, e.g., CAVs, other vehicles, Roadside Units (RSU).

*Autonomous Motion Subsystem (AMS)* computes the Route to destination, produces a Full World Representation using internal and external Basic World Representations and gives commands that drive the CAV to the intended destination.

*Motion Actuation Subsystem (MAS)* provides information on the physical environment, and receives and actuates motion commands in the physical world.

The interaction of the 5 subsystems in depicted in *Figure 3*:



*Figure 3 – The CAV subsystems*

The following high-level workflow illustrates the operation of the CAV envisaged by this docum­ent.

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

*Human-CAV Interaction* authenticates the human, interprets the request and passes a command to the *Autonomous Motion Subsystem*. Later, instructions can be corrected/integrated.

*Autonomous Motion Subsystem*:

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

Computes the Route.

Issues the start command.

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

*CAV to Everything* sends the CAV’s Basic World Representation to and receives Basic World Representations from CAV-compat­ible devices in range. Then itand sends all Basic World Representations to *Aut­onomous Motion Sub­system.*

*Autonomous Motion Subsystem*:

Receives and fues the Basic World Representations to compute the Full World Repres­entation.

Computes a Path.

Issues commands to *Motion Actuation Subsystem* to move the CAV accordingly.

While the CAV moves, humans

Interact and hold conversation with *Human-CAV Interaction* and possibly other humans on board.

Issue commands.

Request views of the environment (Full World Representation) etc.

Interact with (humans in) other CAVs.

## Human-CAV Interaction (HCI)

### Use Case description

Human-CAV Interaction operated based on the principle that the CAV is impersonated by an avatar, selected/produced by the CAV rights-holder. The visible features of the avatar are head face and torso, and the audible feature is speech that embeds as much as possible the sentiment, e.g., emotion, that would be displayed by a human driver.

The CAV’s avatar is reactive to:

1. The Environment, e.g., it can show an altered face because a human driver has done what it considers an improper action.
2. A human, e.g., it shows an appropriate face to a human in the cabins who has made a joke.
3. Etc.

Other forms of interaction are:

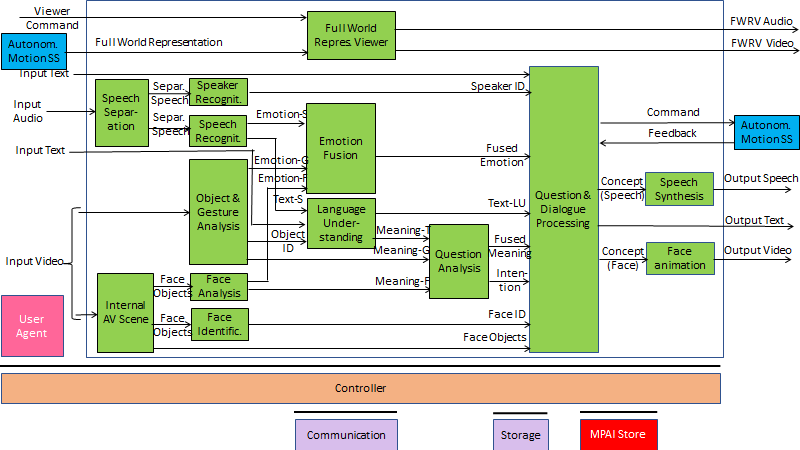
1. CAV authenticates human interacting with it.
2. A human issues commands to a CAV, e.g.:
   1. Commands to Autonomous Motion Subsystem, e.g.: go to a Waypoint or display Full World Representation (see 5.3), etc.
   2. Other commands, e.g.: turn off air conditioning, turn on radio, call a person, open window or door, search for information etc.
3. A human entertains a dialogue with a CAV, e.g.:
   1. CAV offers a selection of offers to human (e.g., long but safe way, short but likely to have interruptions).
   2. Human requests information, e.g.: time to destination, route conditions, weather at destination etc.
   3. Human entertains a casual conversation.
4. A CAV monitors the passenger cabin, e.g.:
   1. Physical conditions, e.g.: temperature level, media being played, sound level, noise level, anomalous noise, etc.
   2. Passenger data, e.g.: number of passengers, ID, estimated age, destination of passengers.
   3. Passenger activity, e.g.: level of passenger activity, level of passenger-generated sound, level of passenger movement, emotion on face of passengers.
   4. Passenger-to-passenger dialogue, two passengers shake hands, or passengers hold everyday conversation.

It is important to point out that, regardless of the fact that vehicles can exhibit different levels of autonomy, the exhibited autonomy should always be adjustable [1]. The system should recognise people as intelligent agents it should inform and be informed by. A CAV should be able to change its level of autonomy to one of several levels during its operation. Such an adjustment may be initiated by a human, another system, or the CAV itself. An important benefit of adjustable, user-centered autonomy is increased user acceptance of the system [39].

### Reference architecture

*Figure 4* represents the Human-CAV Interaction (HCI) Reference Model. The following is noted:

1. A combination of Conversation with Emotion and Multimodal Question Answering AIMs with gesture recognition capabilities covers most Human-CAV Interaction needs.
2. Additional AIMs can be added should new HCI interactions be required.



*Figure 4 – Human-CAV Interaction Reference Model*

The speech of the human is separated from the audio captured from the environment; speaker and speech are recognised; meaning and emotion are extracted from speech; gesture and object information are extracted from video; human is identified, and emotion and meaning extracted from video of the face; emotions are fused; meanings are fused and intention derived to produce speech and face of the avatar interacting with humans; position of humans is computed to provide a realistic gazing. Additionally, human commands and responses from Autonomous Motion Subsystem are processed; Full World Representation is presented to let humans get a complete view of the Environment.

Depending on the technology used (data processing or AI), the AIMs in *Figure 4* may need to access external information, such as Knowledge Bases, to perform their functions. While not represented in *Figure 4*, they will be identified, if required, in the AI Modules subsection.

### Input and output data

*Table 2* gives the input/output data of Human-CAV Interaction.

*Table 2 – I/O data of* *Human-CAV Interaction*

|  |  |  |
| --- | --- | --- |
| **Input data** | **From** | **Comment** |
| Audio | User Outdoor | User authentication  User command |
| Text | User Outdoor | User authentication  User command |
| Text | Passenger Cabin | Social life of user  Commands or interaction with CAV |
| Audio | Passenger Cabin | User’s social life  Commands or interaction with CAV |
| Video | Passenger Cabin | Social life of user  Commands or interaction with CAV |
| Full World Representation | Autonomous Motion SS | For processing by FWR Viewer |
| **Output data** | **To** | **Comments** |
| Text | Autonomous Motion SS | Commands to be executed |
| Synthetic Speech | Passenger Cabin | CAV’s response to passengers |
| Synthetic Face | Passenger Cabin | CAV’s response to passengers |
| Full World Representation | Passenger Cabin | For passengers to view external world |

### AI Modules

*Table 3* gives the AI Modules of the Human-CAV Interaction depicted in *Figure 4*.

*Table 3* *– AI Modules of* *Human-CAV interaction*

|  |  |
| --- | --- |
| **AIM** | **Function** |
| **Speech detection and separation** | 1. Separates relevant speech vs non-speech signals 2. Detects request for dialogue. |
| **Speaker identification** | Recognises speaker. |
| **Speech recognition** | 1. Analyses the speech input. 2. Generates text and emotion output. |
| **Object and gesture analysis** | 1. Analyses video to identify object. 2. Produces the ID of the object in focus. 3. Analyses video. 4. Produces motion and mean­ing of gesture. |
| **Face recognition** | 1. Analyses the video of the face of a human. 2. Recognises the human’s identity. |
| **Face analysis** | 1. Analyses the video of the face of a human. 2. Extracts emotion and meaning. |
| **Language understanding** | 1. Analyses natural language expressed as text using a language model (embedded in AIM). 2. Produces the meaning of the text. 3. Identifies Object ID. |
| **Emotion recognition** | Produces Final Emotion by fusing Emotions from Speech, Face and Gesture. |
| **Question analysis** | 1. Fuses Meanings of Speech, Face and Gesture 2. Analyses the meaning of the sentence. 3. Determines the Intention. 4. Outputs Final Meaning. |
| **Question & dialog processing** | 1. Receives Speaker ID and Face ID. 2. If speaker ID and face ID match, then:    1. Produces a command to Autonomous Motion SS    2. Analyses user’s emotion, intention, meaning and/or ques­tion, text.    3. Produces Concept (speech) and Concept (face). 3. Else, responds appropriately. |
| **Speech synthesis** | Converts Concept (Speech) to Output Speech. |
| **Face animation** | Converts Concept (Face) to Output Video. |
| **Full World Representation Viewer** | 1. Receives Full World Representation (FWR) 2. Presents a FWR view as instructed by human via FWR Com­mands. |

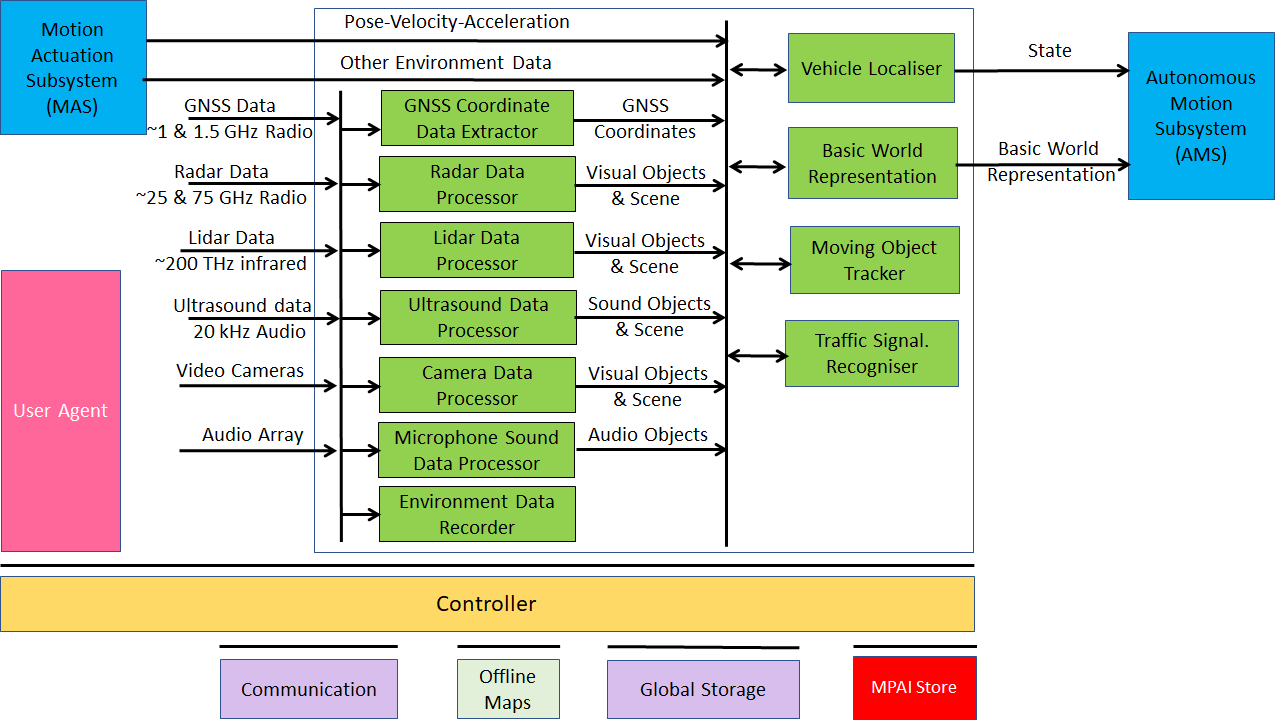
## Environment Sensing Subsystem (ESS)

### Use Case description

The purpose of the ESS is to acquire all sorts of electromagntic and acoustic data directly from its sensors and other physical data of the Environment (e.g., temperature, pressure, humidity etc.) and of the CAV (Pose, Velocity, Accelaration) from Motion Actuation Subsystem with the main goal of creating the Basic World Representation.

### Reference architecture

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



*Figure 5 – Environment Sensing Subsystem Reference Model*

The typical series of operations carried out by the Environment Sensing Subsystem (ESS) is given below. The sequential description of steps does not imply that an action is only carried out after the preceding one has been completed.

1. The CAV gets its Pose and other environment data from:
   1. Global Navigation Satellite System (GNSS).
   2. Vehicle Localiser in ESS.
   3. Other Environment data (e.g., weather, air pressure etc.).
   4. Coordinates data (Pose, Velocity, Acceleration).
2. The CAV creates a Basic World Representation (BWR) by:
   1. Acquiring available Offline maps of its current Pose.
   2. Fusing Visual, Lidar, Radar and Ultrasound data.
   3. Updating the Offline maps with
      1. Other static objects.
      2. All moving objects.
      3. All traffic signalisation.
3. The CAV compresses and stores a subset of the sensor data on board the ESS.

### Input and output data

*Table 4* gives the input/output data of Environment Sensing Subsystem.

*Table 4 – I/O data of* *Environment Sensing Subsystem*

|  |  |  |
| --- | --- | --- |
| **Input data** | **From** | **Comment** |
| Pose-Velocity-Acceleration | Motion Actuation Subsystem | To be fused with GNSS data |
| Other Environment Data | Motion Actuation Subsystem | Temperature etc. to be added to Basic World Representation |
| Global Navigation Satellite System (GNSS) | ~1 & 1.5 GHz Radio | Get Pose from GNSS |
| Radio Detection and Ranging (RADAR) | ~25 & 75 GHz Radio | Get RADAR view of Environment |
| Light Detection and Ranging (LIDAR) | ~200 THz infrared | Get LiDAR view of Environment |
| Ultrasound | 20 kHz Audio | Get 20 kHz view of Environment |
| Cameras (2/D and 3D) | Video (400-800 THz) | Get visible view of Environment |
| Microphones | 16 Hz-16 kHz sound | Get Audible view of Environment |
| **Output data** | **To** | **Comment** |
| State | Autonomous Motion Subsystem | For Route, Path and Trajectory |
| Basic World Representation | Autonomous Motion Subsystem | Locate CAV in Environment |

### AI Modules

*Table 5* gives the AI Modules of Environment Sensing Subsystem.

*Table 5 – AI Modules of* *Environment Sensing Subsystem*

|  |  |
| --- | --- |
| **AIM** | **Function** |
| **GNSS Data Coordinate Extractor** | Computes global coordinates of CAV. |
| **Radar Data Processor** | Extracts electromagnetic scene and objects. |
| **Lidar Data Processor** | Extracts electromagnetic scene and objects. |
| **Ultrasound Data Processor** | Extracts ultrasound scene and objects. |
| **Camera Data Processor** | Extracts visual scene and objects. |
| **Environment Sound Data Processor** | Extracts audible audio scene and objects. |
| **Environment Data Recorder** | Compresses/records a subset of data produced by CAV sensors at a given time. |
| **Vehicle Localiser** | Estimates the current CAV State in the Offline Maps. |
| **Moving Objects Tracker** | Detects, tracks and represents position and velocity of Environment moving objects. |
| **Traffic Signalisation Recogniser** | Detects and recognises traffic signs to enable the CAV to correctly move in conformance with traffic rules. |
| **Basic World Representation Fusion** | Creates Basic World-Representation by fusing Offline Map, moving and traffic objects, and other sensor data |

## CAV-to-Everything (V2X)

### Use Case description

A CAV exchanges information via radio with other entities, e.g., CAVs in range and other CAV-like communicating devices such as Roadside Units and Traffic Lights, thereby improving its Environment perception capabilities. Multicast mode is typically used for heavy data types (e.g., Basic World Representation). Unicast mode may be used in other cases.

CAVs in range and other CAV-like communicating devices install a Communication AIM in the CAV-to-Everything AIW of a CAV. The installed Communication AIM informs the Controller that it has a recognised role classified in term of priority as:

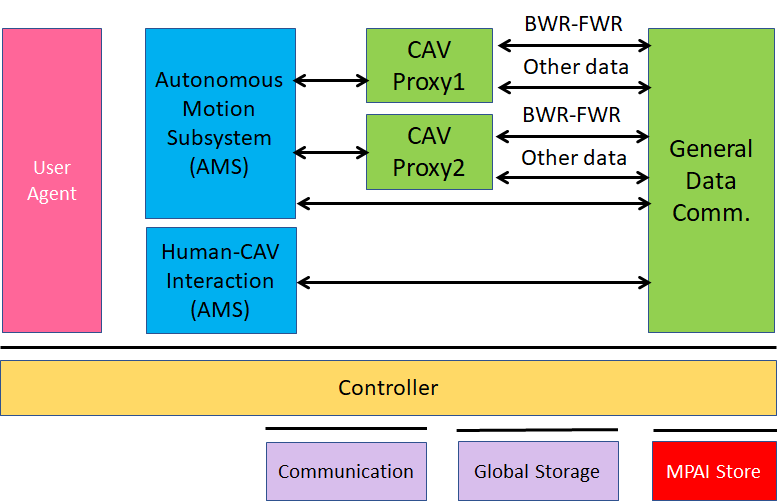
1. Traffic light
2. Fire Truck
3. Police
4. Ambulance
5. Flock Leader.

### Reference architecture

CAVs in range are important not just as sources of valuable information, but also because, by commun­icating with them, each CAV can minimise interfence with other CAVs while pursuing its own goals.

The selected way to achieve this is by having a “CAV Proxy AIM” running in the CAV-to-Everything AIW. The “General Data Communication” AIM is in charge of communication with all communicating devices in range.

The CAV-to-Everything Subsystem Reference Model is given by *Figure 6*.



*Figure 6 – The CAV-to-Everything Subsystem Reference Model*

### Input and output data

#### CAVs within range

*Table 6* gives the data types a CAV broadcasts to CAVs in range.

*Table 6 – I/O data of* *CAV-to-Everything*

|  |  |  |
| --- | --- | --- |
| **Input Data** | **From** | **Comments** |
| Basic World Representation | Other CAVs | A digital representation of the Environment created by ESS. |
| CAV Identity | Other CAVs | Digital equivalent of today’s plate number. Includes Manufacturer and Model information. |
| CAV Intention | Other CAVs | The Path and other motion data relevant to other CAVs |
| Full World Representation | Other CAVs | A digital representation of the Environment created by fusing all available Basic World Representations. |
| Information Messages | Other CAVs | Typical messages a CAV can broadcast. Sources of messages potentially important for CAVs are given by [37, 38]   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 World Representation | Other CAVs | Same as input for all other input data. |
| Full World Representation | Other CAVs | A digital representation of the Environment obtained from the fusion of all available Basic World Representations. |

#### CAV-aware equipment

Such equipment are traffic lights, roadside units, vehicles with CAV communication capabilities.

1. Identity and coordinates (exact coordinate reference).

Recorded Full World Representation (no moving objects) regularly updated via download (may be part of the Offline Map).

Current objects in Environment.

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

Lane markings.

Speed limits.

Pedestrian crosswalks

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

Etc.

These equipment can have one of the following functions:

1. Act as any other CAV in range.
2. Authority to organise motion of CAVs in range.

#### Other non-CAV vehicles

Other vehicles can be scooters, motorcycles, bicycles, other non-CAV vehicles, possibly transmitting their position as derived from GNSS. No response capability is expected. Vehicle 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.

### AI Modules

*Table 11* gives the AI Modules of Autonomous Motion Subsystem.

*Table 7 – AI Modules of CAV-To-Everything Subsystem*

|  |  |
| --- | --- |
| **AIM** | **Function** |
| **General Data Communication** | Communicates with non-CAV sources. |
| **CAV Proxy** | Forwards data of  Local CAV’s AIMs to the remote CAV it represents.  Remote CAV’s AIMs to the appropriate local CAV’s AIMs. |

## 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. Dialogue may follow.
2. Computes the Route satisfying the human’s request.
3. Receives the current Basic World Represen­tation from Environment Sensing Subsystem.
4. While moving, CAV
   1. Transmits the Basic World Representation and other data to CAV-to-Everything.
   2. Receives Basic World Representations and other data from CAV-to-Everything.
   3. Produces the Full World Representation by fusing its own Basic World Representation with those from other CAVs in range.
   4. Plans a Path connecting Poses.
   5. Selects behaviour to reach intermediate Goals acting on information about the Goals other CAVs in range intend to reach.
   6. Defines a Trajectory that
      1. Complies with general traffic rules and local traffic regulations
      2. Preserves passengers’ comfort.
   7. Refines Trajectory to avoid obstacles.
   8. Sends the Motion Actuation Subsystem Commands to take the CAV to the next Goal.
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 [9], are possible depending on the amount and level of available func­tionalities.

### Reference architecture

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

A human activates the CAV requesting to be transported to a waypoint. This activates the Route Planner and the Path Planner which requests the Full World Representation to Full World Representation Fusion which receives and fuses the Basic World Representations from its own and other CAVs’ Environment Sensing Subsystems. The chain Behaviour Selection-Motion Planner-Obstacle Avoider eventually generates a command which is sent to Motion Actuation Subsystem. The decisions of the said chain are recorded.

Diagram

Description automatically generated

*Figure 7 – Autonomous Motion Subsystem Reference Model*

### 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** |
| Human Command | Human-CAV Interaction | Human commands, e.g., “take me home” |
| Basic World Representation | Environment Sensing Subsystem | CAV’s Environment representation. |
| Other Environment Data | Environment Sensing Subsystem | E.g., temperature, air pressure. |
| Other V2X Data | CAV To Everything | Other CAVs and vehicles, and roadside units. |
| Command Feedback | Motion Actuation Subsystem | CAV’s response to Command |
| **Output data** | **To** | **Comment** |
| AMS Response | Human-CAV Interaction | MAS’s response to AMS Command |
| AMS Command | Motion Actuation Subsystem | Macro-instructions, e.g., “in 5s assume a given State”. |
| Full World Representation | CAV To Everything | 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 Pose. |
| **Path Planner** | Generates a set of Paths, considering   1. Current Route. 2. State. 3. Full World-Representation. 4. Traffic Rules. |
| **Behaviour Selector** | Sets a Goal with a Driving Behaviour, to be reached within the Decision Horizon time frame. |
| **Motion Planner** | Defines a Trajectory, from the current State to the current Goal fol­lowing the Behaviour Selector’s Path to the extent possible, satisfying the CAV’s kinematic and dynamic constraints, and considering passengers’ comfort. |
| **Obstacle Avoider** | Defines a new Trajectory to avoid obstacles. |
| **Command** | Instructs the CAV to execute the Trajectory considering the Environment conditions. |
| **Full World-Representation Fusion** | Creates an internal representation of the Environment by fusing infor­mation from itself, CAVs in range and other transmitting units.. |

## Motion Actuation Subsystem (MAS)

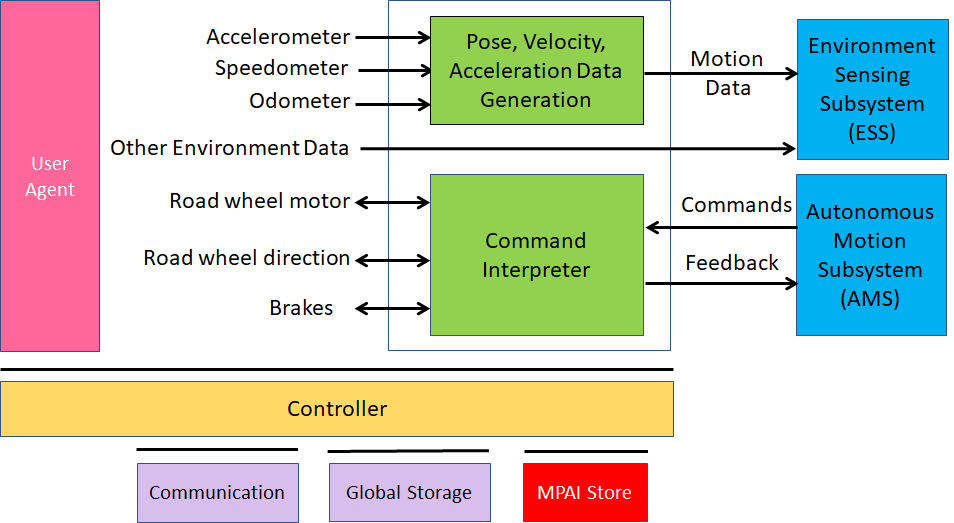
### Use Case description

The Motion Actuation Subsystem:

1. Transmits information gathered from its sensors and its mechanical subsystems to Environ­ment Sensing Subsystem.
2. Receives instructions from Autonomous Motion Subsystem.
3. Translates such instructions into specific commands to its own mechanical subsystems, e.g., road wheels, wheel motors, brakes.
4. Receives feedbacks from its mechanical subsystems.
5. Packages feedbacks into high-level information.
6. Send high-levelinformation to Autonomous Motion Subsystem.

### Reference architecture

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



*Figure 8 – Motion Actuation Subsystem Reference Model*

### Input and output data

*Table 10* gives the input/output data of Motion Actuation Subsystem.

*Table 10 – 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. |
| Road Wheel Motor | Forces road wheels rotation, gives feedback. |
| Road Wheel Direction | Moves road wheels by an angle, gives feedback. |
| Brakes | Acts on brakes, gives feedback. |
| AMS Commands | High-level motion command. |
| **Output** | **Comments** |
| Motion data | Position, velocity, acceleration. |
| Other data | Other environment data. |
| MAS Feedback | Feedback from Command Converter during and after Command ex­ecution |

### AI Modules

*Table 11* gives the AI Modules of Autonomous Motion Subsystem.

*Table 11 – AI Modules of Motion Actuation Subsystem*

|  |  |
| --- | --- |
| **AIM** | **Function** |
| **Pose-Velocity-Acceleration Data Generation** | Transforms odometer, speedometer, accelerometer data to standard data format. |
| **Command Interpreter** | Converts AMS Commands into specific actuation commands to Road wheel motor, Road wheel direction and Brakes. Forwards MAS feedbacks to AMS. |

# Functional Requirements

Functional Requirements developed in this document refer to individual technologies identified as necessary to implement MPAI-CAV Use Cases using AIMs operating in an MPAI AI Framework (AIF). They adhere to the following guidelines:

AIMs are defined to allow implementations by multiple technologies (AI, ML, DP).

DP-based AIMs may need interfaces, e.g., to a Know­ledge Base. AI-based AIM will typically require a learning process, however, support for this process is not included in the document. MPAI may develop further requirements covering that process in a future document.

AIMs can be aggregated in larger AIMs. Consequently, some data flows of aggregated AIMs may no longer be accessible.

MPAI has issued Calls for Technologies for the MPAI-MMC [3] and MPAI-CAE [4] standards and acquired a set of first-generation technologies related to some of the data types listed below. MPAI is ready to consider new technologies related to the data Formats requested in this Chapter if:

1. They support new requirements and/or to enhance capabilities.
2. The need to support such new enhanced capability requirements are documented.

## Human-CAV Interaction

### I/O Data summary

For each AIM (1st column), *Table 12* gives the input (2nd column) and the output data (3rd column).

*Table 12 – I/O data of Human-CAV Interaction AIMs*

|  |  |  |
| --- | --- | --- |
| **AIM** | **Input Data** | **Output Data** |
| **Speech Separation** | Input Audio | Separated Speech |
| **Internal AV Scene** | Input Video | Face Objects |
| **Speaker Recognition** | Separated Speech | Speaker ID |
| **Speech Recognition** | Separated Speech | Emotion (Speech)  Text (Speech) |
| **Object and Gesture Analysis** | Input Video | Object ID  Emotion (Gesture)  Meaning (Gesture) |
| **Face Analysis** | Face Objects | Emotion (Face)  Meaning (Face) |
| **Face Identification** | Face Objects | Face ID |
| **Full World Representation Viewer** | Full World Representation  Viewer Command | FWRV Audio  FWRV Video |
| **Emotion Fusion** | Emotion (Speech)  Emotion (Face)  Emotion (Gesture) | Fused Emotion |
| **Language Understanding** | Text (Speech)  Input Text  Object ID | Text (Language Understanding)  Meaning (Text) |
| **Question analysis** | Meaning (Text)  Meaning (Gesture)  Meaning (Face) | Fused Meaning  Intention |
| **Question and dialogue processing** | Input Text  Speaker ID  Fused emotion  Text (Speech)  Fused Meaning  Intention  Face ID  Face Objects | Command/Request  Feedback/Response  Concept (Speech)  Output Text  Concept (Face) |
| **Speech synthesis** | Concept (Speech) | Output Speech |
| **Face animation** | Concept (face) | Output Video |

### Audio

Monochannel Audio is the digital representation of an analogue audio signal sampled at a frequency between 8-192 kHz with a number of bits between 8 bits/sample and 32 bits/sample and a quantisation that is linear or companded.

**To respondents**

Respondents are invited to comment on these definitions and/or provide specific restrictions suitable to CAV-HCI.

**To respondents**

Respondents are requested to propose a coded representation of the above commands coordinated with the requirements of the with the Autonomous Motion Subsystem Responses. Proposals of coded representation of additional commands are welcome.

### Verbal Interaction

Some commands given to the Autonomous Motion Subsystem are:

1. Go to a Waypoint.
2. How long does it take to get there.
3. Park close to a Waypoint.
4. Drive faster.
5. Drive slowly.
6. Display Full World Representation.

Some of the responses of the Autonomous Motion Subsystem are:

1. Enumeration of possible routes with major features of each route.
2. Enumeration of possible parking places with major features of each place.
3. Announcement of obstacles preventing the expeditious accomplishment of the Command.
4. Announcement that the desired Waypoint has been reached.

**To respondents**

Respondents are requested to propose a coded representation of the above Commands/Responses. Proposals of coded representation of additional responses are welcome.

### Concept Expression (Face)

MPAI-MMC [3] specifies a Lips Animation format.

**To Respondents**

In this call, MPAI is looking for a technology that can animate head and face of the avatar with the purpose to represent:

1. Motion of head when speaking.
2. Motion of face muscles and eyeballs.
3. Turning of gaze to a particular person.
4. Emotion of the associated spoken sentence.
5. Meaning of the associated spoken sentence.

### Concept Expression (Speech)

MPAI-MMC [3] specifies Text With Emotion as Reply (speech) format.

**To Respondents**

Respondents are requested to propose a “Concept to Speech” format with the following requir­ements:

1. Capability to represent Emotions varying in time in the synthesised Speech.
2. Capability to represent Meanings varying in time in the CAV reply.

### Emotion

MPAI-MMC [3] specifies an extensible 3-level Basic Emotion Set.

**To respondents**

Respondents are requested to comment on the suitability of the technology standardised in [3] for the purpose of supporting human dialogue with a CAV. In case this is considered unsuitable, respondents are requested to motivate their assessment and provide an extension of the MPAI Basic Emotion Set or a new solution.

### Face identity

The Face Identity shall be able to represent the identity of a limited number of faces.

**To respondents**

Respondents are requested to propose a face identification system suitable for a limited number of faces.

Proposals of a face identification usable in the context of a company renting CAVs to customers are welcome.

### Face Objects

In order for the HCI Subsystem to have a full understanding of what is happening in the passenger cabin (e.g., to have a more natural audio-visual interaction with the passengers, recording of what happened in the cabin etc.), the HCI Subsystem needs to represent the data acqu­ired from the cabin. The current use is:

1. To extract the face of a passenger for the purpose of extracting Emotion and Identity.
2. To determine the exact location of a passenger in the cabin in order to animate the CAV’s Avatar Face in such a way that the Avatar gazes into the eyes of the passenger it is talking to.

**To respondents**

Respondents are invited to propose a Face Objects format satisfying the above requirements to be used as input to Face Analysis, Face Identification and Question and Dialogue Processing.

### Full World Representation

The Full World Representation requirements are developed in the context of Autonomous Motion Subsystem requirements.

**To respondents**

Respondents are invited to comment.

### Full World Representation Viewer commands

The requirements of FWR interaction will be developed once the FWR requirements are defined.

**To respondents**

Respondents are invited to comment.

### Intention

MPAI-MMC [3] specifies a digital representation format for Intention.

**To respondents**

Respondents are requested to comment on the suitability of the technology standardised in [3] for CAV purposes.

### Meaning

MPAI-MMC [3] specifies a digital representation format for Meaning.

**To respondents**

Respondents are requested to comment on the suitability of the technology standardised in [3] for CAV purposes.

### Object Identifier

MPAI-MMC [3] specifies a digital representation format for Object Identifier to be used to identify objects held in the hand of a person.

**To respondents**

Respondents are requested to comment on the suitability of the technology standardised in [3] for CAV purposes.

### Speaker Identity

The current Speaker Identity requirements demand the ability to identify a limited number of Speakers.

**To respondents**

Respondents are requested to propose a Speaker Identification methods suitable for a limited number of speakers.

Proposals of a Speaker Identification method usable in a content of a company renting CAVs to customers are welcome.

### Text

MPAI-MMC [3] specifies ISO/IEC 10646, Information technology – Universal Coded Character Set (UCS) [5] as digital Text representation to support most languages in use.

**To respondents**

Respondents are invited to comment on this choice.

### Video

Video is intended for use in the passenger cabin. MPAI-MMC [3] specifies Video as:

1. Pixel shape: square
2. Bit depth: 8 or 10 bits/pixel
3. Aspect ratio: 4/3 or 16/9
4. 640 < # of horizontal pixels < 1920
5. 480 < # of vertical pixels < 1080
6. Frame frequency 50-120 Hz
7. Scanning: progressive
8. Colorimetry: ITU-R BT709 or BT2020
9. Colour format: RGB or YUV
10. Compression, either:
    1. Uncompressed;
    2. Compressed according to one of the following standards: MPEG-4 AVC [6], MPEG-H HEVC [7], MPEG-5 EVC [8]

**To respondents**

Respondents are invited to comment on MPAI’s choice for 2D Video.

Respondents are also requested to propose a data format for an array of cameras having video+depth as the baseline format or other 3D Video data formats.

## Environment Sensing Subsystem

### I/O Data summary

For each AIM (1st column), *Table 13* 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 13 – Environment Sensing Subsystem data*

|  |  |  |
| --- | --- | --- |
| **AIM or Subsystem** | **Input** | **Output** |
| **Vehicle Localiser** | GNSS Coordinates | State |
| Pose-Velocity-Acceleration |
| Offline Maps |
| **Environment Recorder** | State | -- |
| Basic World Representation |
| Other Environment Data |
| **GNSS Coordinate Data Extractor** | GNSS data | Global coordinates |
| **Radar Data Processor** | Radar data | Visual Objects and Scene |
| **Lidar Data Processor** | Lidar data | Visual Objects and Scene |
| **Ultrasound Data Processor** | Ultrasound data | Visual Objects and Scene |
| **Camera Data Processor** | Camera data | Visual Objects and Scene |
| **Microphone Sound Data Processor** | Microphone data | Sound Objects and Scene |
| **Traffic Signalisation Detector** | Visual Objects and Scene | Traffic signals  Traffic rules |
| **Moving Objects Tracker** | Visual Objects and Scene | Moving objects’ states |
| **Basic World Representation Fusion** | State | Basic World Representation |
| Offline maps |
| Visual Objects and Scenes |
| Static and moving objects |
| Traffic signals |

### Microphone Array Audio

Microphones are used to capture the external sound, (e.g., for noise suppression inside the passen­ger cabin, but also to add the sound dimension to the Full World Representation by using the Audio from the Environment captured by the external Microphone Array.

MPAI-CAE specifies Interleaved Multichannel Audio [4].

**To Respondents**

Respondents are requested to comment on the usability of the specified technology for MPAI-CAV and/or propose an Audio Array Format suitable to create a 3D sound field representation of the Environment to be added to the Basic World Representation and used inside the passenger cabin, e.g., to cancel Environment noise.

### Audio Objects

The sound field of the Environment is captured by the external Microphone Array Audio, and Objects are extracted and added to the Basic and eventually Full World Representation after receiving information from other CAVs in range.

**To Respondents**

Respondents are requested to propose an Audio Objects Format that provides information about audio objects identified in the Environment with their semantics and the degree of accuracy with which objects have been represented.

### Basic World Representation

Data from different information sources, e.g., CAV’s Environment sensors and Offline maps are combined to one comprehensive Basic World Representation (BWR) [31]. The BWR ensures that all CAV functions base their decisions on the same knowledge base, thus ensuring consistency of system operation.

The requirements of the BWR are:

1. All perceived objects that impact the path decision process in the Decision Horizon Time shall be represented in the BWR
2. Each object in the BWR shall be described by
   1. Its ID.
   2. Its State.
   3. Its physical characteristics, e.g., static or dynamic.
   4. Its bounding box (as a minimum) and its full shape if known.
   5. Its semantics (e.g., other CAVs or other objects).
   6. An accuracy estimate.
3. The ground (roads etc.) shall be described with all traffic signalisations, including roads and lane geometry, topology, and lane-specific traffic rules.
4. The BWR shall have the ability to scale as to the level of structuredness of the Environment increases.
5. The BWR shall have a scalable representation that allows fast access to critical data.
6. The BWR shall include the Audio Objects of 6.2.3.

**To Respondents**

Respondents are requested to propose a Basic World Representation data format satisfying the requirements. Proposals with justified extended requirements will be considered.

### GNSS Coordinates

**To Respondents**

Respondents are requested to provide a format for the coordinates and the accuracy of the data.

### GNSS Data

Global Navigation Satellite Systems (GNSS) provide spatial information with different accuracies. GNSS can only be relied on when reception conditions are above a certain level. This excludes GNSS in tunnels or urban canyons.

Some data formats are:

1. GPS Exchange Format (GPX) provides an XML schema providing a common GPS data format that can be used to describe waypoints, tracks, and routes.
2. World Geodetic System (WGS) includes the definition of the coordinate system's fundamental and derived constants, the ellipsoidal (normal) Earth Gravitational Model (EGM), a description of the associated World Magnetic Model (WMM), and a current list of local datum transfor­mations.
3. International GNSS Service (IGS) SSR is a format used to disseminate real-time products to support the IGS (igs.org) Real-Time Service. The messages support multi-GNSS and include corrections for orbits, clocks, DCBs, phase-biases and ionospheric delays. Extensions are planned to also cover satellite attitude, phase centre offsets and variations and group delay variations.

**To Respondents**

Respondents are requested to propose a single GNSS data format that is capable to represent the features of all GNSS types.

### Lidar Data

Radio Detection and Ranging (RADAR), LiDAR and Ultrasound are active sensors based on “time-of-flight”, i.e., they measure distance and speed based on the time it takes for a signal to hit an object and be reflected back.

Unlike Radar, however, it operates in the µm range – ultraviolet, visible, or near infrared light. It sends an electromagnetic signal and receives the reflected signal back. These are the features of a typical eye-safe LiDAR:

1. Has a frequency of ~200 THz and a wavelength ~1.5 µm (the visible range is 0.4 to 0.75 µm).
2. Measures the range in each pixel (called also voxels).
3. Pixel grayscale is measured by the intensity variation of the reflected light.
4. The colour of an object can be measured by using more than one wavelength.
5. Velocity can be measured using the Doppler shift in frequency due to motion, or by measuring the position at different times.
6. Micro-motion can be measured using the Doppler shift measured with a coherent LiDAR.
7. Produces 100 kpoints/frame or 1.35 Mbytes: 32\*3 bits (coordinates) +16 bits (ref­lectance). Today 200 kpoints/frame are reasonable.
8. Angular resolution is 0.1º and the vertical field is 40º.
9. A Lidar scan captured at 25 fps generates 270 Mbit/s, i.e., 33.75 Mbytes/s.

The LAS (LASer) format is a binary file format for LiDAR point cloud data specified by the American Society for Photogrammetry and Remote Sensing (ASPRS) [23].

Pcap isa well-established data format for Lidar scans [24, 25, 26]. Other formats are listed in [28]. E57 is one of them.

**To Respondents**

Respondents are invited to provide a LiDAR data format that facilitates identification, tracking and digital representation of objects to produce Visual Objects and Scene (Lidar) as required by 6.2.16.

### Moving Object Tracker Data

Moving Object Tracker receives the Visual Objects and Scene data from the different sources – Lidar, Radar, Cameras, Ultrasound, Environment Sound – and provides a list of Visual Objects where each Object has the following associated data:

1. Spatial coordinates
2. Bounding Boxes
3. Coordinated of vertices of Bounding Boxes
4. Velocity and Acceleration
5. Accuracy of the data provided.

**To Respondents**

Respondents are requested to propose a format for the Objects and their list that is friendly to the Basic World Representation format.

### Offline Maps

An Offline Map or HD maps or 3D maps is a roadmap with cm-level accuracy and a high environ­mental fidelity reporting the positions of pedestrian crossings, traffic lights/signs, barriers etc. at the time the Offline Map has been created.

Worth noting are:

1. Navigation Data Standards [30] calls itself “The worldwide standard for map data in autom­ot­ive eco-systems”. Their NDS specification covers data model, storage format, interfaces, and protocols.
2. SharedStreets [34] Referencing System is a global non-proprietary system for describing streets.

**To Respondents**

Respondents are requested to propose an Offline Map Format. The Format should support different levels of conformance.

### Radar Data

Radar operates in the mm range. It can detect vehicles (CAVs and trucks) because they typically reflect Radar signals while objects that are smaller and have less reflectance, e.g., pedestrians and motorcycles have a poor reflectance. In a busy environment, the reflections of big vehicles can overcome a motorcycle’s and a child next to a vehicle can go undetected, while a can may produce an image out of proportion to its size.

The main features of Radar are:

1. Measures distance.
2. Is independent of environment.
3. ImaLow resolution (objects detected, not classified).
4. Short range radar in the 25 GHz band, distance is computed.
5. Long range radar in the 76-77 GHz, detects objects and measures speed @ ≤ 250 m. Typical ranges of long-range radar (LRR) systems are 80-200 m. The antenna is small because the wavelength is ~3.5-4 mm. Atmospheric absorption limits interference with other systems. A multitask 94-GHz pulse Doppler radar has 25-cm radial and 1.5 degrees angular resolution

Radar sensors build a representation of the environment based on the observation of complex, scattered radio waves, from which information of an object’s distance and velocity can be derived.

Known Radar data formats include [27]:

1. OPERA BUFR format (Paulitsch et al., 2010).
2. hdf5 formats (Michelson et al., 2011).
3. NetCDF files generated by the commercial EDGE software.
4. hdf5 files generated by the commercial GAMIC software.
5. German Weather Services quantitative local scan format (DX).
6. Quantitative composite format (RADOLAN, see German Weather Service, 2004).

**To Respondents**

Respondents are invited to propose a format of Radar images that facilitates identification, tracking and representation of objects to produce Visual Objects and Scene (Radar) as required by 6.2.17.

### State

State is the set of the following CAV attributes at a given time:

1. Pose, Velocity and Acceleration
2. Orientation, Angular Velocity and Angular Acceleration.

**To Respondents**

Respondents are requested to propose a State Format suitable for use in CAVs.

### Traffic Signalisation

Traffic Signalisation types are:

1. Traffic signs
2. Road signs
3. Placement signs
4. Acoustic signs
5. Traffic lights

**To Respondents**

Respondents are requested to propose a set of Traffic Signalisation Descriptors.

### Ultrasound Data

These are the main features of Ultrasound:

1. Operates at 20 kHz.
2. Is independent of environment.
3. Images have low resolution.
4. Works on a limited range (≤ 10 m)

**To Respondents**

The Ultrasound File Format initiative has defined the Ultrasound File Format (UFF) format [22].

Respondents are invited to propose an Ultrasound Format that facilitates identification, tracking and representation of sound objects to produce Visual Objects and Scene (Radar) as required by 6.2.18.

### Video Camera data

The expected output is a scene representation used by Moving Object Tracker, Traffic Signalisation Recogniser and Basic World Representation.

**To Respondents**

Respondents are invited to provide a data Format for RGB-D cameras.

### Visual Objects and Scene (Camera)

The expected output is a scene representation used by Moving Object Tracker, Traffic Signalisation Recogniser and Basic World Representation.

**To Respondents**

Respondents are invited to provide a Format for scenes captured by cameras. The format should be sufficiently generic to be capable to be used for – or adapted for use in – scenes captured by Radar, Lidar and Ultrasound devices.

### Visual Objects and Scene (Lidar)

The expected output is a scene representation used by Moving Object Tracker, Traffic Signalisation Recogniser and Basic World Representation.

**To Respondents**

Respondents are invited to provide a Format for scenes captured by Lidars. The format should be sufficiently generic to be capable to be used for – or adapted for use in – scenes scenes captured by Radar, Video and Ultrasound devices.

### Visual Objects and Scene (Radar)

The expected output is a scene representation used by Moving Object Tracker, Traffic Signalisation Recogniser and Basic World Representation.

**To Respondents**

Respondents are invited to provide a Format for scenes captured by Radars. The format should be sufficiently generic to be capable to be used used for – or adapted for use in – scenes captured by Lidar, Video and Ultrasound devices.

### Ultrasound Objects and Scene (Ultrasound)

The expected output is a scene representation used by Moving Object Tracker, Traffic Signalisation Recogniser and Basic World Representation.

**To Respondents**

Respondents are invited to provide a Format for scenes captured by Ultrasound. The format should be sufficiently generic to be capable to be used for – or adapted for use in – scenes scenes captured by Lidar, Radar and Video devices.

## CAV to Everything

### Summary of CAV to Everything data

*Table 15* gives, for each AIM (1st column), the input data (2nd column) and the output data (3rd column).

*Table 14 –CAV to Everything data*

|  |  |  |
| --- | --- | --- |
| **CAV AIM** | **Input** | **Output** |
| General Data Communication | CAV identity and model | CAV identity and model |
| General Data Communication | State-Path-Trajectory | State-Path-Trajectory |
| General Data Communication | Basic World Representation | Basic World Representation |
| General Data Communication | Full World Representation | Full World Representation |
| General Data Communication | Messages | Messages |
| General Data Communication | Basic World Representation | Basic World Representation |
| CAV Proxy | Data from General Data Communication  Data from AIMs | Data to General Data Communication  Data to AIMs |
| AMS | Basic World Representation | Basic World Representation |
| HCI | Data from General Data Communication | Data to General Data Communication |

### Basic World Representation

As in Environment Sensing Subsystem.

**To Respondents**

No response requested here. Comments welcome.

### CAV Identifier

The CAV identification system should carry 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

**To Respondents**

MPAI requests proposals for universal CAV identification system. Justified proposals for inclus­ion of additional data in the CAV Identifier are welcome

### Events

Events is used to provide CAV with information that is useful for its travel.

Examples are:

1. Road blocked at waypoint x,y,z
2. Traffic jam at waypoint x,y,z
3. ...

**To Respondents**

MPAI requests proposals for events, their semantics and coded representation.

### Full World Representation

Defined in Autonomous Motion Subsystem.

**To Respondents**

No response requested here. Comments welcome.

### Path

Defined in Autonomous Motion Subsystem.

**To Respondents**

No response requested here. Comments welcome.

### State

Defined in Autonomous Motion Subsystem.

**To Respondents**

No response requested here. Comments welcome.

### Trajectory

Defined in Autonomous Motion Subsystem.

**To Respondents**

No response requested here. Comments welcome.

## Autonomous Motion Subsystem

### Summary of Autonomous Motion Subsystem data

*Table 15* gives, for each AIM (1st column), the input (2nd column) and the output data (3rd column).

*Table 15 – CAV Autonomous Motion Subsystem data*

|  |  |  |
| --- | --- | --- |
| **CAV/AIM** | **Input** | **Output** |
| **Route Planner** | Pose  Destination | Route  Estimated time |
| **Full World Representation Fusion** | State | Full World Representation |
| Offline Maps |
| Basic World Representations |
| Other Environment Data |
| **Path Planner** | State | Set of Paths |
| Route |
| Traffic Rules |
| **Behaviour Selector** | State | Path |
| Route |
| Full World Representation |
| **Motion planner** | Path | Trajectory |
| **Obstacle Avoider** | Full World Representation Trajectory | Trajectory |
| **Command** | Feedback | Command |

### Basic World Representation

Defined in Environment Sensing Subsystem.

**To Respondents**

No response requested here. Comments welcome.

### Command/Response

Defined in Human-to-CAV subsystem.

**To Respondents**

No response requested here. Comments welcome.

### Full World Representation

The elements of the FWD are:

1. Appropriate portion of the offline map.
2. Physics of the environment: weather, temperature, air pressure, ice and water on the road).
3. For each object: ID, position, velocity, acceleration bounding box (more than a box, if available), semantics, flags (e.g., warning).
4. For CAVs, the Path and bounding box or the shape of the body, if available.
5. Road structure.
6. Local traffic signalisation.
7. Scalable representation such that the following is possible:
   1. Fast access to different data depending on the AIM who needs to access.
   2. Deliberative and reactive actions.
8. The estimated accuracy of each data element based on the CAV’s measurements and the Basic World Representations received from CAVs in range.

### Goal

A particular State.

**To Respondents**

No response requested. Comments welcome.

### Offline map

Defined in Environment Sensing Subsystem.

**To Respondents**

No response requested here. Comments welcome.

### Path

A sequence of Poses in the Offline Map

**To Respondents**

No response requested here. Comments welcome.

### Pose

The position of an oject, including the CAV, in the Environment.

**To Respondents**

A format to represent Pose is requested.

### Route

A sequence of Waypoints.

**To Respondents**

A Route Format compatible with a proposed Offline Map Format is requested.

### State

Defined in Environment Sensing Subsystem.

**To Respondents**

No response requested here. Comments welcome.

### Traffic rules

The traffic rules should be digitally represented to realise a route [33].

**To Respondents**

MPAI requests a digital representation of traffic rules satisfying the following requirements:

1. Produce the traffic rules from a given set of traffic signals
2. Produce the traffic signals from the traffic rules.

A Traffic Ontology is a possible solution.

### Traffic Signals

Format to represent traffic signals on a road and around it.

**To Respondents**

MPAI requests a Traffic Signals Format capable to represent

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

### Trajectory

The Path and the States that allows a CAV to start from a State and reach another State in a given amount of time without violating Traffic Rules and affeccting passengers’ comfort.

**To Respondents**

A digital representation of Trajectory is requested.

### Velocity

Defined in Environment Sensing Subsystem.

**To Respondents**

No response requested here. Comments welcome.

## Motion Actuation Subsystem

### Summary of Motion Actuation Subsystem data

*Table 15* gives, for each AIM (1st column), the input data (2nd column) from which AIM (column) and the output data (3rd column).

*Table 16 –Motion Actuation Subsystem data*

|  |  |  |
| --- | --- | --- |
| **CAV/AIM** | **Input** | **Output** |
| **Command Interpreter** | Command from AMS  Road Wheel Motor Feedback  Road Wheel Direction Feedback  Brakes Feedback | Feedback to AMS  Road Wheel Motor Command  Road Wheel Direction Command  Brakes Command |
| **Pose, Velocity, Acceleration Data Generation** | Accelerometer  Speedometer  Odometer | Motion Data |

### Accelerometer data

An accelerometer is an electronic sensor that measures the acceleration forces acting on a CAV. An accelerometer measures proper acceleration, i.e., the acceleration of a body in its own instantaneous rest frame, not to be confused with coordinate acceleration, i.e., acceleration in a fixed coordinate system. Therefore, an accelerometer at rest on the surface of the Earth measures an acceleration straight upwards of g ≈ 9.81 m/s2. In free fall (falling toward the centre of the Earth at ≈ 9.81 m/s2) measures zero.

**To Respondents**

Respondents are requested to propose a single Accelerometer data format.

### Brakes Command

The result of the interpretation of AMS Command to Brakes.

**To Respondents**

Respondents are requested to propose a set of command messages.

### Brakes Feedback

The feedback of Brakes to Command Interpreter.

**To Respondents**

Respondents are requested to propose a set of feedback messages.

### Command from AMS

The Command issued by the AMS

**To Respondents**

Respondents are requested to propose a set of high-level command messages.

### Feedback to AMS

The Feedback of Command Interpreter summarising the Feedbacks.

**To Respondents**

Respondents are requested to propose a set of high-level feedback messages

### Motion Data

**To Respondents**

Respondents are requested to propose a Motion Data Format bearing in mind that Motion Data will be used to create the CAV State by adding GNSS information.

### Odometer Data

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

**To Respondents**

Respondents are requested to propose a single Odometer Data Format.

### Other Environment Data

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

**To Respondents**

Respondents are requested to propose a set Environment Data Formats.

### Road Wheel Direction Command

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

**To Respondents**

Respondents are requested to propose a set of Road Wheel Direction Commands

### Road Wheel Direction Feedback

The feedback of Road Wheel Direction to Command Interpreter.

**To Respondents**

Respondents are requested to propose a set of Road Wheel Direction Feedbacks

### Road Wheel Motor Command

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

**To Respondents**

Respondents are requested to propose a set of Road Wheel Motor Commands

### Road Wheel Motor Feedback

The feedback of Road Wheel Motor to Command Interpreter.

**To Respondents**

Respondents are requested to propose a set of Road Wheel Motor Feedbacks

### Speedometer

An electronic sensor that measures the instantaneous speed of a CAV.

**To Respondents**

Respondents are requested to propose a single Speedometer data format.

# Data privacy

A 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 World 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 World 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 my 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 Worlds 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 World Representations of all CAVs in range and its own, a CAV can create a pretty 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 Inertial Measurements Unit.

# Annex 1 – General MPAI Terminology (Normative)

The Terms used in this standard whose first letter is capital and are not already included in *Table 1* are defined in *Table 17.*

*Table 17 – 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 Workflow (AIW) | An organised aggregation of AIMs implementing a Use Case receiving AIM-specific Inputs and producing AIM-specific Outputs according to its Function. |
| AI Module (AIM) | A processing element receiving AIM-specific Inputs and producing AIM-specific Outputs according to according to its Function. |
| Application Standard | An MPAI Standard designed to enable a particular application domain. |
| Channel | A connection between an output port of an AIM and an input port of an AIM. The term “connection” is also used as synonymous. |
| Communication | The infrastructure that implements message passing between AIMs |
| Component | One of the 7 AIF elements: Access, Communication, Controller, Internal Storage, Global Storage, 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 Format | The standard digital representation of data. |
| Data Semantics | The meaning of data. |
| 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. |
| 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. |
| Internal Storage | A Component to store data of the individual AIMs. |
| Identifier | A name that uniquely identifies an Implementation. |
| Implementation | 1. An embodiment of the MPAI-AIF Technical Specification, or 2. An AIW or AIM of a particular Level (1-2-3) conforming with a Use Case of an MPAI Applic­ation Standard. |
| Interoperability | The ability to functionally replace an AIM with another AIM having the same Interoperability Level |
| Interoperability Level | The attribute of an AIW and its AIMs to be executable in an AIF Implem­en­tati­on and to:   1. Be proprietary (Level 1) 2. Pass the Conformance Tes­ting (Level 2) of an Applic­ation Standard 3. `Pass the Perform­ance Testing (Level 3) of an Applic­ation Standard. |
| Knowledge Base | Structured and/or unstructured information made accessible to AIMs via MPAI-specified interfaces |
| Message | A sequence of Records transported by Communication through Channels. |
| Normativity | The set of attributes of a technology or a set of technologies specified by the applicable parts of an MPAI standard. |
| Performance | The attribute of an Implementation of being Reliable, Robust, Fair and Replicable. |
| 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 |
| Profile | A particular subset of the technologies used in MPAI-AIF or an AIW of an Application Standard and, where applicable, the classes, other subsets, options and parameters relevant to that subset. |
| Record | A data structure with a specified structure |
| Reference Model | The AIMs and theirs Connections in an AIW. |
| Reference Software | A technically correct software implementation of a Technical Specific­ation containing source code, or source and compiled code. |
| Reliability | The attribute of an Implementation that performs as specified by the Application Standard, profile and version the Implementation refers to, e.g., within the application scope, stated limitations, and for the period of time specified by the Implementer. |
| Replicability | The attribute of an Implementation whose Performance, as Assessed by a Performance Assessor, can be replicated, within an agreed level, by another Performance Assessor. |
| Robustness | The attribute of an Implementation that copes with data outside of the stated application scope with an estimated degree of confidence. |
| Service Provider | An entrepreneur who offers an Implementation as a service (e.g., a recommendation service) to Users. |
| 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 |  |

# Annex 2 - 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.

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.

# Annex 3 – 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.

|  |  |
| --- | --- |
| 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 is a high-level description of the MPAI ecosystem operation applicable to fully conforming MPAI implementations:

1. MPAI establishes and controls the not-for-profit MPAI Store (step 1).
2. MPAI appoints Performance Assessors (step 2).
3. MPAI publishes Standards (step 3).
4. Implementers submit Implementations to Performance Assessors (step 4).
5. If the Implementation Performance is acceptable, Performance Assessors inform Implementers (step 5a) and MPAI Store (step 5b).
6. Implementers submit Implementations to the MPAI Store (step 6); The Store Tests Confor­mance and security of the Implementation.
7. Users download Implementations (step 7).

Text

Description automatically generated with low confidence

*Figure 9 – The MPAI ecosystem operation*

The Ecosystem operation allows for AIW and AIF Implementations to be:

1. Proprietary: security is verified and Conformance to MPAI-AIF Tested (Level 1).
2. Conforming to an MPAI Application Standard: security is verified and Conformance to the relevant MPAI Application Standard Tested (Level 2).
3. Assessed to be Reliable, Robust, Fair and Replicable (Level 3).

and have their Interoperability Level duly displayed in the MPAI Store.

# Annex 4 – Datasets for CAV research

**nuScenes**

The nuScenes dataset (https://nuscenes.org/) is a large-scale autonomous driving dataset with 3d object annotations. It features:

* Full sensor suite (1x LIDAR, 5x RADAR, 6x camera, IMU, GPS)
* 1000 scenes of 20s each
* 1,400,000 camera images
* 390,000 lidar sweeps
* Two diverse cities: Boston and Singapore
* Left versus right hand traffic
* Detailed map information
* 1.4M 3D bounding boxes manually annotated for 23 object classes
* Attributes such as visibility, activity and pose
* New: 1.1B lidar points manually annotated for 32 classes
* New: Explore nuScenes on SiaSearch
* Free to use for non-commercial use

For a commercial license contact nuScenes@motional.com

nuImages is a large-scale autonomous driving dataset with image-level 2d annotations. It features:

* 93k video clips of 6s each (150h of driving)
* 93k annotated and 1.1M un-annotated images
* Two diverse cities: Boston and Singapore
* The same proven sensor suite as in nuScenes
* Images mined for diversity
* 800k annotated foreground objects with 2d bounding boxes and instance masks
* 100k 2d semantic segmentation masks for background classes
* Attributes such as rider, pose, activity, emergency lights and flying
* Free to use for non-commercial use

**Road Hazard data**

Otonomo real-time Road Hazard data from connected passenger vehicles powers diverse road safety use cases, including mapping, accident predictions, smart cities and many more. The Otonomo Vehicle Data Platform secures, cleanses and normalizes the hazard data to make it more valuable and accessible for diverse use cases.

https://info.otonomo.io/hazard-data-datasheet-lp

# Annex 5 – ETSI Technical Report

ETSI specifies the Collective Perception Service (CPS) in its Technical Report [15]. The CPS includes the format and generation rules of the Collective Perception Message (CPM).

The CPM message format is (H=header, C=container, M=mandatory, O=optional).

*Table 18 – ETSI Collective Perception Message format*

|  |  |  |  |
| --- | --- | --- | --- |
| PDU header | H | M | protocol version, message ID and Station ID. |
| Management | C | M | transmitter type (e.g., vehicle or RSU) and position. |
| Station Data | C | O | transmitter heading, velocity, or acceleration etc. |
| Sensor Information | C | O | transmitter (e.g., speed, heading, or acceleration)  capabilities of the vehicle’s sensors. |
| Perceived Object | C | O | detected objects (e.g., distance, speed and dimensions)  time at which the measurements were done.  A CPM can report up to 128 detected objects |
| Free Space Addendum | C | O | free space areas/volume within the sensor detection areas |

Every 0.1s a CPM is generated if one of the 3 conditions is satisfied

no CPM has been generated in the last 1s

a new object has been detected

since last CPM sending info about a previously detected object (it must have an ID)

the following attributes have changed:

Absolute position ΔP > 4 m

Absolute speed ΔV > 0.5 m/s

more than 1s has passed (ΔT > 1 s).

ETSI makes use of a common coordinate system. A vehicle can communicate its absolute coordinates roll, pitch and yaw (Attitude).

Different CPM generation rules have been investigated [16].

# Annex 6 – Some CAV Communication Technologies

The following categories of vehicular communication are part of the literature or industry effort:

|  |  |  |
| --- | --- | --- |
| V2V | Vehicle-to-Vehicle | communication between vehicles to exchange information about the speed and position of surrounding vehicles |
| V2I | Vehicle-to-Infrastructure | communication between vehicles and road infrastructure. |
| V2X | Vehicle-to-Everything | communication between a vehicle and any entity that may affect, or may be affected by, the vehicle |
| V2R | Vehicle-to-Roadside | communication between a vehicle and Road Side Units (RSUs). |
| V2P | Vehicle-to-Pedestrian | communications between a vehicle and (multiple) pedestrian device(s) and to other vulnerable road users, e.g., cyclists, in close proximity |
| V2S | Vehicle-to-Sensors | communication between a vehicle and its sensors on board |
| V2D | Vehicle-to-Device | communication between a vehicle and any electronic device that may be connected to the vehicle itself |
| V2G | Vehicle-to-Grid | communication with the power grid to sell demand response services by either returning electricity to the grid or by throttling their charging rate |
| V2N | Vehicle-to-Network | broadcast and unicast communications between vehicles and the V2X management system and also the V2X AS (Application Server) |
| V2C | Vehicle-to-Cloud | communication with data centres and other devices connected to the internet |

Technologies exist that support at least some aspects of the communication types of the table:

Radio access, e.g., visible light communication, mmWave, Cellular-V2X, and 5G

Radio resource management (RRM) for vehicular communication using cellular technology

3GPP Release 14: air interfaces and core network technologies to support V2X communic­ation.

Vehicular ad hoc network (VANET)

Dedicated Short-Range Communication (DSRC): 5.9 GHz band with a range of ~300 metres.

Software defined vehicular networks (SDVN)

Internet of vehicles (IoV)

Protocol stack of the intelligent transportation system (ITS)

Cooperative Awareness Messages (CAMs) messages related to the status of CAV’s sent via wireless broadcast in VANETs.

Cooperative or collective perception improve CAV’s perception beyond the sensors’ detection range.

Traffic situation can be extracted from Local dynamic map (LDM) that aggregates CAMs.