PRISSMA Project
Plateforme de Recherche et d'Investissement pour la Sûreté
et la Sécurité de la Mobilité Autonome
04/2021 - 04/2024

[L4.4] METHODOLOGY AND PROCEDURE FOR TESTING IN REAL CONDITION

Méthodologie et procédure de mise en œuvre d'intervention

Main authors: Thibaut Jonville (CEREMA), Leo Maisonobe (STRMTG), Omar Hafa (RATP), Youssoufa Tahirou (VALEO), Christophe Bohn (IRT-SystemX), Emmanuel Arbaretier (APSYS-AIRBUS), Marie-Françoise Guyonnaud (ATC France)

Keywords: Real-world test and validation, Criteria, Methods, Infrastructure, POC, Assessment

Abstract. This deliverable is linked to Task 4.3 of the PRISSMA project. The main objectives of this task and this deliverable are to build the methodology of assessment, in adequacy between the description of the pathways (in particular the Paris2Connect POC) and the Valeo and RATP POCs proposed as part of the project in real conditions. The first part reminds us of the main route descriptors (in relation to the work of Task 8), the second part shows us an example applied to a section of the Paris2Connect pathway and the third part makes the link with the Valeo and RATP POCs.

Résumé. Ce livrable est lié à la Tâche 4.3 du projet PRISSMA. Les principaux objectifs de cette tâche et de ce livrable sont de construire la méthodologie d'évaluation d'un VA, en adéquation entre la description des parcours (notamment le POC Paris2Connect) et les POC Valeo et RATP proposés dans le cadre du projet en conditions réelles. La première partie nous rappelle les principaux descripteurs du parcours (en lien avec les travaux de la Tâche 8), la seconde partie nous montre un exemple appliqué sur une section du parcours Paris2Connect et la troisième partie fait le lien avec les POC Valeo et RATP.

Table des matières

Definitions and Acronyms

Definitions:

Dynamic Driving Task (DDT)	All operational and tactical functions performed in real time required for vehicle movement, including: 1) control of the lateral and longitudinal movement of the vehicle, 2) monitoring of the road environment 3) reaction to events in the roadway environment, 4) preparation and reporting of manoeuvres, 5) activation of visibility functions. NOTE: Excluded are strategic functions such as trip scheduling, definition of times and positions of departure and arrival points.
Operational Domain (OD)	Real-world conditions that an ADS may experience [1] Set of operating conditions, including, but not limited to, environmental, geographical, and time-of-day restrictions, and/or the requisite presence or absence of certain traffic or roadway characteristics [2]
Operational Design Domain (ODD)	Operating conditions under which an ADS is designed to operate safely [2] Specific conditions under which a given driving automation system is designed to function [3]
Object and Event Detection and Response (OEDR)	DDT subtasks that include monitoring the driving environment (detecting, recognizing and classifying objects and events), as well as executing an appropriate response to such objects and events.
Planned event	A situation that is known in advance, for example, at the time of activation, such as a crossing point (e.g., highway exit, etc.) and requires a transition request.
Unexpected event	A situation that is not known in advance, but which is assumed to be highly likely to occur, e.g., road work, inclement weather, approach of an emergency vehicle, lack of lane markings, falling truck load (collision), and which requires a transition request.
EGO	Name used for automated/autonomous vehicle. In the analysis of a driving situation involving multiple vehicles, the EGO vehicle is the subject of the study, the one whose behaviour we seek to understand or control in interactions with the other vehicles involved in this situation

Acronyms

ADS: Automated Driving System
AI: Artificial Intelligence
ALKS: Automated Lane Keeping System
ARTS: Automated Road Transport System
CAM: Cooperative Awareness Message
CPM: Collective Perception Message
DDT: Dynamic Driving Task
DENM: Decentralized Environmental Notification Message
GRVA: Working Party on Automated/Autonomous and Connected Vehicles (at UNECE)
IVIM: Infrastructure to Vehicle Information Messages
KPI: Key Performances Indicators
LOM: Loi d'Orientation des Mobilités (French Law : Mobility Orientation Law)
MAPEM: Map-data Messages
MRC: Minimal Risk Condition
MRM: Minimal Risk Manoeuver
NHTSA: US National Highway Traffic Safety Administration
POC: Proof of Concept
ODD: Operational Design Domain
OEDR: Object and Event Detection and Response
POI: Point of interest
RSU: Road Side Unit
SPATEM: Signal Phase and Timing Messages
UNECE: United Nations Economic Commission for Europe
VRU: Vulnerable Road User

1. Introduction

Purpose of the L4.4 procedure is to establish a testing and driving methodology in real conditions, with objective to build a strategy for homologation. For that, L4.4 used all resources available to find the right guidance to evaluate behaviour of a vehicle with IA, and through a specific itinerary.
This procedure WP4 (in real conditions) has to be implemented in correlation with WP2 (simulation) & WP3 (controlled environment).
This procedure is describing the method developed and used around available cases to complete WP4 and especially the final L4.5 test & driving.
The method is organized to collect a maximum of information and data in order to establish a concrete evaluation.

It means that the scenarios will be chosen to get all chances to cross-check a maximum of data and to analyse vehicle behaviour, in order to build the main guidelines for a global evaluation for approval and qualification.

From the perspective of EU ADS and vehicle homologation at level 4, such a procedure currently does not cover everything (vehicle and onboard interfaces only). In the PRISSMA project, the proposed procedure so far is solely focused on vehicle evaluation around EU ADS scenarios and KPIs. This lays the groundwork for a broader future ambition to envision a new demonstration phase encompassing the entirety of a fully connected vehicle system with infrastructure and route elements (traffic intersections, Spatem/Mapem, communication with infrastructure, DENM...).

2. Context

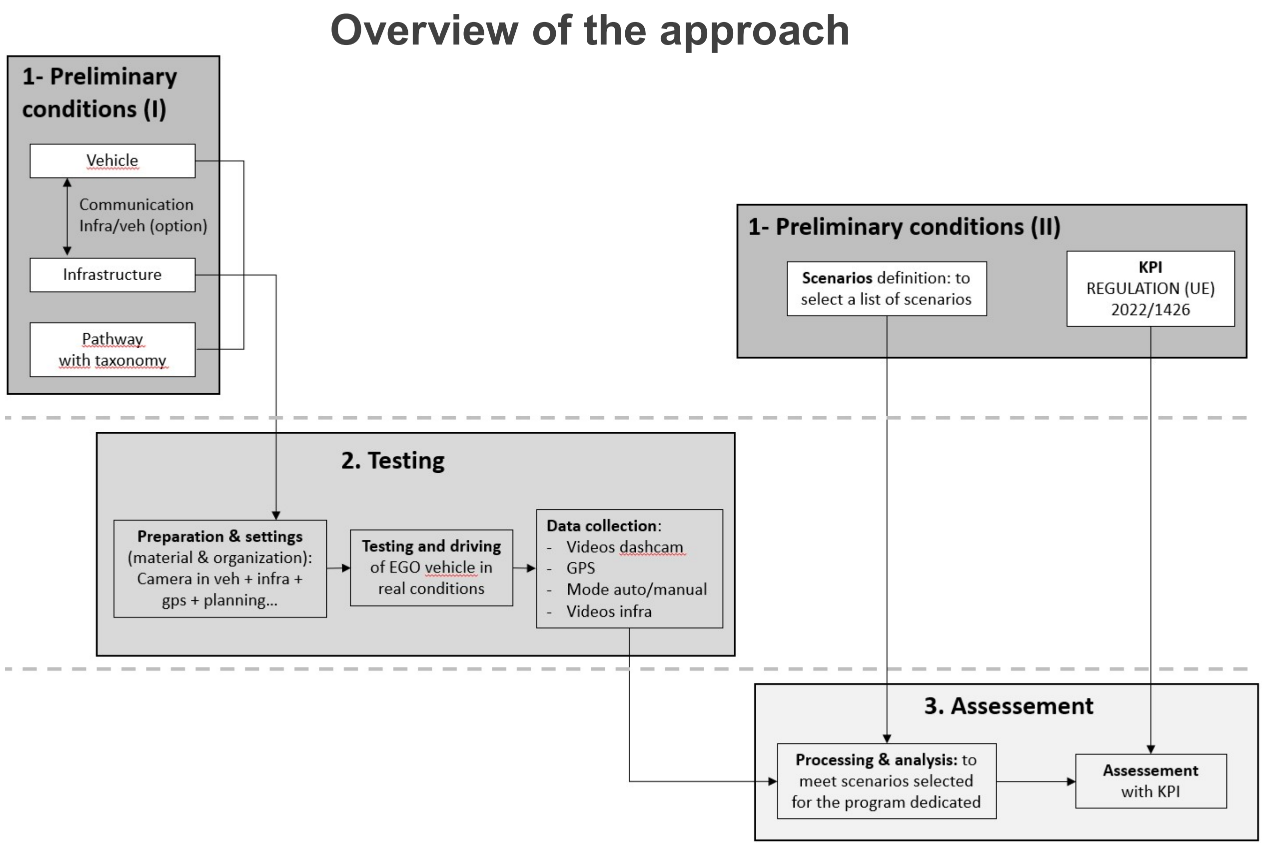
The evaluation context will encompass all circumstances, information, and elements surrounding the experiment. It will include various environmental factors relevant to understanding and interpreting the situation. The context will provide the framework for grasping the meaning and significance of encountered events, actions, or interactions. This will notably encompass elements such as location, background, and external factors (environment) influencing the situation and resulting decisions. Understanding the context is essential for thorough analysis and proper interpretation of events in the evaluation method for automated and autonomous vehicles within a territory.

Within this contextual paragraph, everything laying the groundwork for the application of the procedure will be addressed, notably:

- Experimentation objectives: Defining the objectives and expectations of the experimentation. This may include aspects related to safety, performance, user experience, etc.
- Experimentation environment: Identifying the location and conditions under which the experimentation will be conducted. This could be an urban area, a highway, specially designed lanes, or a mixture of different types of roads.

3. Methodology

Before delving into the details of the methodology, here is a diagram of the overall approach summarizing the process: from prerequisites to evaluation.

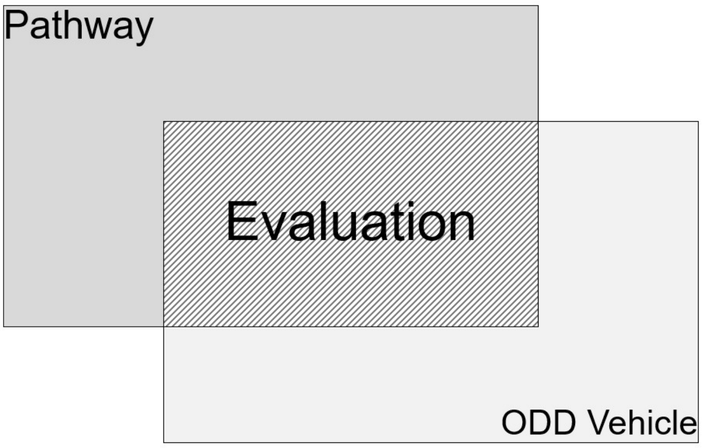


3.1. Preliminary conditions

First of all, safety criteria must be the top priority in real-world testing; hence, safety is the primary prerequisite that will be emphasized throughout WP4. Safety is closely linked to the technology deployed and its state-of-the-art, making it crucial to specify the versions of tools and software used.

The evaluation will focus on the vehicle's ability to operate within the defined route and environment. Therefore, in its final version of EGO evaluation experimentation, the vehicle's Operational Design Domain (ODD) must align with the defined route. It is anticipated that a portion of the route may lie outside the vehicle's ODD, and conversely, that some aspects of the vehicle's ODD may lie outside the route domain, indicating that the EGO's capabilities may exceed or be partially adapted to the defined route. This underscores the importance of selecting and defining the route, as it determines the expectations and scope of the evaluation.

Thus, the aim is to define the EGO's capabilities within a known and broad framework, considering the most extensive use cases to cover a wide range of encountered situations, in order to evaluate the EGO's level by identifying its limits and considering areas where its capabilities exceed the scope of the evaluation.



The prerequisites necessary for the implementation of the methodology (Test setup, data collection, analysis, etc.) will be listed as questions classified in the table below.

These prerequisites are identified by importance to facilitate the implementation of the methodology: "Mandatory," "Recommended," or "Unimportant."

The list of questions should be completed and "calibrated" according to the intended purposes in the evaluation process for homologation.

Questionnary preliminary conditions	Mandatory	Recommended	Unimportante
Vehicle			
Is there a safety-driver? (cf. LOM regulation)	X		
Is it a level 4 designed vehicle?			
Is the timestamp (gps) calibrated with the infrastructure and other devices?		X	
Is veh. ODD adapted with pathway?	X		
What softwares are used in the system?	X		
What are data sources? Videos, CAM, GPS...			
- Videos	X		
- CAM		X	
- GPS: csv			X
Infrastructure			
Is the timestamp (gps) calibrated with the vehicle and other devices?			
What softwares are used?			
What are data sources? Videos, Spatem/Mapem...			
- Videos		X	
- Spatem/Mapem		X	
Pathway			
Is the pathway detailed through a taxonomy?	X		
Does the pathway cover the entire expected assessment?	X		
Scenarios			
Are scenarios are described in a support?	X		
How many scenarios have been defined?			
Does the list of scenarios cover the entire expected assessment?			
KPI & metrics			
Are metrics from EU regulation 2022/1426?	X		
Does metrics cover the entire expected assessment?	X		
Other			
What are weather conditions?	X		

3.1.1. Vehicle level SAE 4

The methodology could be applied to a level 3 vehicle. However, within the PRISSMA project, the methodology requires a level 4 vehicle.

The manufacturer will provide the associated homologation dossier.

The manufacturer will provide the list of equipment, devices, and supervision systems comprising the vehicle (Lidar, cameras, etc.).

In real-world conditions, the vehicle must ensure the safety of the environment in which it operates.

Note: Within the PRISSMA experimentation, the objective is not to evaluate the vehicle participating in the experiment, but to develop and test an evaluation method. Indeed, the elements of PRISSMA are components provided and made available by partners within a consortium, rather than components specifically developed for this evaluation for individual homologation purposes.

3.1.2. Infrastructure

In the perspective of a broader and more comprehensive evaluation, an infrastructure will complement the setup within the evaluation framework.

The entity responsible for the infrastructure will provide the list of equipment, devices, and supervision systems comprising the infrastructure (Lidar, cameras, etc.).

This infrastructure serves as the bridge between the vehicle POC and the supervision POC. As defined in the document, the infrastructure will provide an augmented view to the vehicle navigating the road. It can alert to the risk of traffic congestion on the road due to crowd movements, with a certain concentration of pedestrians in specific areas.

3.1.3. Pathway

The pathway must be fully characterized and sequenced according to the taxonomy defined in deliverable L8.11: infrastructure (road type, road edge, geometry, construction zones, vertical and horizontal signage), scenes (specific areas), environmental conditions (lighting), traffic conditions (density, users, safety), operational requirements (speeds, maneuvers) ...

The definition and quality of the evaluation framework will also depend on the level of characterization of the route and the necessary elements identified and affecting the vehicle systems or detected by the infrastructure.

3.1.4. Scenarios

A selection of scenarios to be played in real conditions will be realized for demonstration of AI interaction/decision.

This list of scenarios will need to be established and constructed around the chosen pathway. The selected scenarios must be consistent with REGULATION (EU) 2022/1426. This regulation provides a number of quantitative and qualitative metrics for evaluating the behavior of a given vehicle in a defined situation.

Therefore, it is necessary to have the ability to find, within the chosen pathway (using a taxonomy proposed in L8.11 within the framework of PRISSMA, and all proposed descriptors that must be as reliable and relevant as possible), the opportunity to evaluate metrics through concrete scenarios grouped under families of functional and logical scenarios.

Similarly, as a prerequisite, the evaluation will be based on defined scenarios that fall within the vehicle's ODD. The expectation is that the vehicle responds to what it is designed for.

	ODD vehicle	Out of ODD
Scenarios defined	EVALUATION	Out of evaluation
Out of scenario	Out of evaluation	Out of evaluation

This holds true within the scope of an experimentation aimed at building the methodology.

In the case of an evaluation for homologation purposes, however, this list of scenarios must be constructed to address the evaluation of a vehicle system or subsystems or functions within the defined framework. Consequently, these scenarios can establish a common evaluation framework for assessing vehicles from different manufacturers that must meet a common functional requirement (see WP6).

This evaluation framework is highly strategic as it entails validating a vehicle within a specific context, a given technology, and a precise configuration, thereby committing to a technical state through a prism of knowledge available at a given time.

These defined scenarios are not intended to evaluate what may not arise during experimentation or behaviors that emerge subsequently in similar configurations. To illustrate, if, for instance, the EGO can replicate and validate scenario “X” 10 times during an evaluation, the evaluation would no longer cover the vehicle if the scenario is not replicated once it's put into operation.

3.1.5. KPI & metrics

The KPIs and metrics are derived from REGULATION (EU) 2022/1426. Thus, for the evaluation, we consider both quantitative and qualitative metrics.
As mentioned above, these metrics are associated with scenarios identified on the defined route. Any changes in regulations could render these metrics or parts thereof inadequate, or even obsolete.

3.2. *Test implementation*

3.2.1. Preparatory phase & test plan

3.2.1.1. Onboard equipment (original equipment)

The onboard equipment in the vehicle needs to be identified and described. The manufacturer of the vehicle being evaluated will provide the list of equipment characterizing the vehicle in its automated mode.

3.2.1.2. Additional equipment (dashcam, gps…)

In the context of an evaluation, additional equipment may be installed on board the vehicle to complement the evaluation. The additional equipment will be identified and selected.

3.2.2. Real conditions driving phase

The vehicle will operate on a selected and defined route. Driving conditions must be defined based on the vehicle's ODD and the functional capabilities specified by the manufacturer.

In real conditions, the primary criterion is the safety of the environment or the ODD in which the vehicle will operate.

3.2.3. Data: compilation and exploitation

All vehicle and infrastructure equipment must facilitate data collection for the purpose of evaluating the vehicle's behavior in its environment. This volume of data should serve as the database for compilation, utilization, and analysis.

3.3. Assessment implementation

3.3.1. Vocabulary

For the remainder of the document, it is important to define the vocabulary used.

- **Stop Time:** The duration during which the vehicle is stationary (speed equal to 0 km/h) and no interaction occurs in autonomous mode. For example, during stops at bus stations, traffic lights, pedestrian crossings, and yielding to the right. This time is excluded from the temporal analysis of driving times.
- **Driving Time:** The duration of automated vehicle circulation in the observed area excluding stop times.
- **Incident Concepts**

  There are 2 types of incidents:
  - **"Stop":** The automated vehicle stops in the middle of the road (excluding stop times mentioned above);
  - **"Slowdown":** The automated vehicle significantly slows down in the middle of the road (speed less than or equal to 5 km/h in urban areas), with its cruising speed averaging 30 km/h.

  These incidents may have several possible causes, which are analyzed later in this document:
  - Vulnerable User Cause: All vulnerable road users, such as pedestrians, cyclists, electric scooter riders, etc.;
  - Motorized User Cause: All motorized road users, such as motorcycles, passenger cars, trucks, etc.;
  - User Cause: A collective term comprising both vulnerable and motorized road users;
  - Without Apparent Reason: No evidence around the vehicle identifies the cause;
  - Other: Causes such as a bird, a branch on the road, an unintentional modification of the infrastructure, etc….

3.3.2. Review of video sequences from dashcams

The video sequences from the dashcams are initially sorted by the vehicle's driving period and then manually reviewed by operators. During this process, two tables are completed by the operators. The first table concerns the situations encountered in the vicinity of the AV (Automated Vehicle) as well as the general context of the observations. The second table details the observed AV incidents and their respective causes.

The next phase involves verifying the consistency of the encountered situations with the selected scenarios.

Finally, with the scenarios identified within the situations, it is possible to proceed to the evaluation phase according to the criteria identified in REGULATION (EU) 2022/1426.

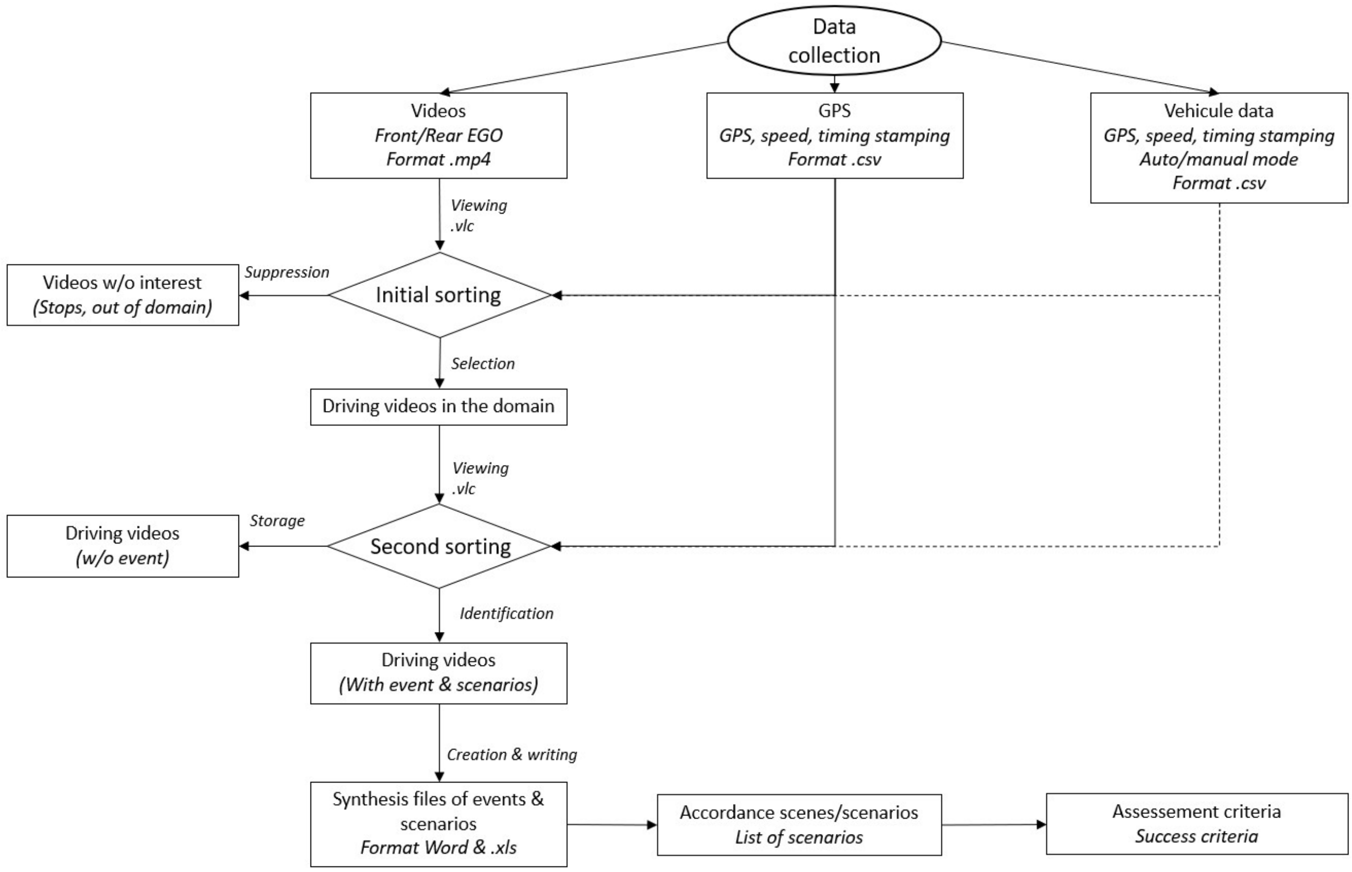


Figure 1: Videos analysis approach & assessment

➢ **<u>First sorting: selection of driving phases within the domain</u>**

The initial sorting aims to reduce the amount of data to be analyzed and focus on specific data that may reveal situations to be evaluated.

- **Second sorting: recording and searching for events during driving**

A sorting table is completed regarding the events.
The following are recorded in the sorting table:

- Stop times in seconds are subtracted from the video times to obtain the actual driving times of the AV;
- Contextual information: Weather (sunny, cloudy, or rainy), maximum AV speed on the route, and passage times.

Date & Video N°	Sheet N°	Weather	Time	Speed	Section	EGO position	Traffic on front	Traffic on rear	VRU	Observations
20231016_xxxxxx	1	Sunny	10:48:43	34 km/h	20	Head	Clear	Crowded	Pedestrian	EGO slowdown to let pedestrian on right

Figure 2 Logsheet of 'occurred events'

It details, for each 'AV stop' incident, its cause, the involved user(s), and the duration of immobilization. For slowdowns, the involved user(s) and the cause are noted. An observation and description of the situation are recorded.

Particular attention is paid to the position and trajectory of each considered user (-5m from the AV). During analysis, the type of user and their mobility are identified: stationary, moving, or crossing the EGO's trajectory (from right to left or from left to right).

The types of users are: Vulnerable Road User (UVR), Pedestrian, Bicycle (including Electric Personal Transportation Devices 'EDP'), Motorcycle (2RM), Passenger Car (VL), and Heavy Goods Vehicle (PL);

- **Compilation of observed scenarios and evaluation**

Finally, the next phase involves compiling the observed scenarios to proceed with the evaluation of the vehicle's behavior in the environment.

3.3.3. Review of infrastructure video sequences

The video sequences from infrastructure shots provide primarily information on traffic signal intersections, UVR/EGO interactions, and traffic/EGO interactions at these locations.

During this analysis, infrastructure images allow for observing and better understanding complex situations where only vehicle images would be too limited. Infrastructure enables observing traffic context (smooth flow, congestion, etc.), observed behaviors and incidents, vehicles or users affected by the EGO's speed when it starts or passes without stopping, and also allows observing situations before and after the EGO's passage event

3.3.4. Video data analysis

Once the sorting tables are duly completed, the following metrics can be compiled:

- Type of encountered scenario;
- Relevant metric;
- Duration of observation in the studied area: Time of videos analyzed in the area;
- EGO stop time (not counted);
- EGO driving time and speed in the area;
- Number and duration of incidents (Stops, Slowdowns);
- Users encountered by the EGO by user type and trajectories;
- Distribution of incident causes by trajectories and by users;
- Number of vehicles affected by the EGO's speed.

Scenario N°	**Data**	**Sheet N°**	**Metrics**													
			Qualitative					**Quantitative**								
			1	2	3	4	5	1	2	3	4	5	6	7	8	9
1	Yes/No	5			✓						O					
2																
3																
4																
5																
6																
7																
8																
9																
10																
11																
12																
13																
14																
15																

✓ Passed   O Unexpected

Table 10: Assessment matrix

The results will be compiled in the form of a matrix. Each detected event will be identified and assigned to the dedicated scenario(s). Then, each event will be evaluated according to the corresponding metrics.

In the end, the evaluation matrix will help identify trends in vehicle behavior. The granularity of the evaluation will depend on the quantity of events recorded to determine recurrences and repeatabilities.

Conclusion

In this deliverable L4.4, we have described the procedure implemented within the PRISSMA project. The purpose of this procedure is to lay the groundwork for a methodology aimed at evaluating systems or products integrating one or more AI.

We have adopted a 'bottom-up' approach, which was dictated by the project. Indeed, in order to carry out this procedure in real conditions and conduct an experimentation to collect data and be able to analyze them, we have used the latest means known to date and available from the consortium partners. We have benefited from an existing route dedicated to automated vehicle experimentation, equipped with infrastructure, and then we have had access to a level 4 vehicle.

In terms of regulation, it should be noted that the EU ADS in the context of homologating a level 4 vehicle does not cover everything, but only the vehicle and the onboard interfaces.

In the PRISSMA project, the procedure proposed to date is focused solely on vehicle evaluation around predefined scenarios and a selection of KPIs from the EU ADS.

This represents a foundation for future ambition to envision a new phase of demonstration encompassing an entire system of a fully connected vehicle with infrastructure and the route (traffic intersections, Spatem/Mapem, infrastructure communication, DENM...).

REFERENCES

[1] Commission Implementing Regulation (EU) 2022/1426 of 5 August 2022 laying down rules for the application of Regulation (EU) 2019/2144 of the European Parliament and of the Council as regards uniform procedures and technical specifications for the type-approval of the automated driving system (ADS) of fully automated vehicles - Official Journal of the European Union - 26.8.2022