Submitted by GRVA

Informal document **WP.29-189-07**
189th WP.29, 7-9 March 2023
Provisional agenda item 2.3.
(Based on GRVA-15-26)

New Assessment/Test Method for Automated Driving (NATM) Guidelines for Validating Automated Driving System (ADS) – amendments to ECE/TRANS/WP.29/2022/58

I. Background

1. At the 178th (June 2019) session of the United Nations Economic Commission for Europe (UNECE)'s World Forum for Harmonization of Vehicle Regulations (WP.29), Terms of Reference (ToRs) (WP.29/1147/Annex VI) for the Informal Working Group on Validation Methods for Automated Driving (VMAD) were developed. VMAD's mandate under these ToRs is to develop assessment methods, including scenarios, to validate the safety of automated systems based on a multi-pillar approach including audit, simulation/virtual testing, test track, and real-world testing. Throughout this document, safety encompasses the safe performance of automated driving systems and System Safety.

2. Also at the 178th session, WP29 adopted the Framework document on automated/autonomous vehicles (WP.29/2019/34/Rev.2) herein referred to as the Framework document. The Framework document instructed VMAD to develop a 'new assessment/test method for automated driving' (NATM) for consideration during the 183rd (March 2021) session of WP.29.

3. To inform this work, VMAD developed the NATM master document which outlines a conceptual framework for validating the safety of automated driving systems. The first version of this document was adopted at the 184th session (June 2021) of WP29 (ECE/TRANS/WP.29/1159). The second version was submitted to the 12th session (January 2022) of GRVA (ECE-TRANS-WP29-GRVA-2022-02e).

4. Building on this conceptual work, VMAD was instructed by WP29 (ECE/TRANS/WP.29/1159) to undertake the development of NATM guidelines that could provide direction to developers and contracting parties of the 1958 and the 1998 UN vehicle regulations agreements on recommended procedures for validating the safety of automated driving systems (ADS).

II. Purpose and scope

5. This guidelines document represents current best practices identified by the Informal Working Group on Validation Methods for Automated Driving (VMAD) for validating the safety of automated driving systems (ADS) using the NATM. These guidelines aim to provide clear direction for validating the safety of an ADS in a manner that is repeatable, objective and evidence-based, while remaining technology neutral and flexible enough to foster ongoing innovation by the automotive industry. The intended audience for these guidelines includes both developers of ADS technologies as well as contracting parties to both the 1958 and the 1998 UN vehicle regulations agreements.

6. Validating ADS safety is a highly complex task which cannot be done comprehensively nor effectively through one validation methodology alone. As a result, it is recommended to adopt a multi-pillar approach for the validation of ADS, composed of a scenarios catalogue and five validation methodologies (pillars):

(a) Simulation/virtual testing,

(b) Track testing

(c) Real world testing

(d) Audit/assessment

(e) In-service monitoring and reporting

7. The following chapters of this guidance document explore each of these NATM components in further detail and outline a number of recommendations and consideration when using them to validate ADS safety. Further information on how the components of the NATM guidelines (i.e., the scenarios catalogue and pillars) operate together to produce an efficient, comprehensive, and cohesive process is discussed at the end of the document.

8. ADS technology is continuously evolving. This document will continue to evolve and be regularly updated from the outcomes of future research and testing as well as through the work of WP.29 working groups.

9. In particular, updates to these guidelines will take into consideration the deliverables from the informal working group on Functional Requirements for Automated Vehicles (FRAV), which has been tasked by WP.29 to develop safety performance requirements, including measurable/verifiable criteria, to assess ADS safety.

10. Subject to direction from GRVA and WP.29, once the guidelines have reached a sufficient state of maturity it is anticipated that this document will be used to help inform the development of regulatory requirements that meet the needs of both 1958 and 1998 Agreement parties (subject to approval by WP.29).

III. Definitions

11. The introduction of ADS and related technologies has resulted in a proliferation of new terms and concepts. To ensure consistency, a glossary of terms and definitions used in the NATM guidelines are attached in Annex I. These terms, which are used throughout the document, have been italicized for reference. This glossary will be further developed and updated on an ongoing basis. Where applicable, VMAD will ensure these terms are consistent with those adopted by WP.29, GRVA, and other GRVA Informal Working Groups, including definitions agreed upon by FRAV.

IV. Applying a multi-pillar approach to the NATM

12. As previously noted, validating ADS safety is a highly complex task which cannot be done comprehensively nor effectively through one validation methodology alone. As a result, it is recommended to adopt a multi-pillar approach for the validation of ADS. This approach is comprised of the scenarios catalogue and five validation methodologies (pillars).

13. The multi-pillar approach and scenarios catalogue are described below and are explored in greater detail in subsequent sections of this document:

(a) A scenario catalogue, consisting of descriptions of ~~real-world~~ driving situations that may occur during a given trip and is a tool used by the NATM-pillars to systematically validate the safety of an ADS;

(b) Simulation/virtual testing which uses different types of simulation toolchains to assess the compliance of an ADS with the safety requirements on a wide range of virtual scenarios including some which would be extremely difficult if not impossible to test in real-world settings. The credibility of the simulation/virtual testing is included in this topic;

(c) Track testing using a closed-access testing ground with various scenario elements to test the capabilities and functioning of an ADS;

(d) Real world testing using public roads to test and evaluate the performance of ADS related to its capacity to drive in real traffic conditions;

(e) Audit/assessment procedures which establish how manufacturers will be required to demonstrate to safety authorities the capabilities of their ADS. This iwill be based on the evidence from their documentation, simulation, test-track, and/or real-world testing of the ADS. The audit will validate that hazards and risks relevant to the ADS have been identified and that a robust and consistent safety methodology has been implemented including safety-by-design. The audit will also verify that robust processes/mechanisms/strategies (i.e., safety management system) are in place to ensure the ADS meets the relevant safety requirements throughout the vehicle lifecycle. It shall also assess the complementarity between the different pillars of the assessment and the overall scenario coverage;

(f) In-service monitoring and reporting addresses the in-service safety of the ADS after its placing on the market. It relies on the collection of fleet data in the field to assess whether the ADS continues to be safe when operated on the road. This data collection can also provide information to help develop new scenarios or variations of existing scenarios for the scenarios catalogue allowing the whole ADS community to learn from major ADS accidents/incidents.

V. Scenarios catalogue

14. At this relatively early stage in the development of ADS, much of the existing literature that assesses the current state of ADS development uses metrics such as miles/kilometres travelled in real-world test situations with the absence of a collision, a legal infraction, or a disengagement by the vehicle's ADS.

15. Metrics such as kilometres travelled without a collision, legal infraction, or disengagement can be helpful for informing public dialogue about the general progress being made to develop ADS. Such measurements on their own, however, do not provide sufficient evidence to the international regulatory community that an ADS will be able to safely navigate the vast array of different situations a vehicle could reasonably be expected to encounter.

16. Furthermore, validation through real world testing alone would be time and cost prohibitive, potentially requiring an ADS to drive billions of kilometres without incident to prove that it has significantly better safety performance than a human driver. It would also not be feasible to replicate this testing later if there was a change to the system that needed to be re-validated.

17. With these considerations in mind, it is recommended that a scenarios-based approach be used to systematically organize safety validation activities in an efficient, objective, repeatable, and scalable manner.

18. Scenarios based validation consists of reproducing specific situations that exercise and challenge the capabilities of an ADS-equipped vehicle to operate safely.

19. Going forward, VMAD will establish a catalogue of scenarios that can be used by the various NATM pillars to validate the functional safety requirements established by FRAV.

A. What is a traffic scenario?

20. A scenario is a description of one or more driving situations that may occur during a given trip[1]. Scenarios can involve many elements, such as roadway layout, types of road users, objects exhibiting static or diverse dynamic behaviours, and diverse environmental conditions (among other factors).

B. Ensuring adequate scenario coverage

21. It is recommended that the scenarios-based validation methods include adequate coverage of relevant, nominal, failure, critical, and *complex scenarios* to effectively validate an ADS. To note: "Coverage" refers to the degree to which scenarios sufficiently incorporates driving situations in order to validate the relevant requirements defined by FRAV. Sufficient coverage is essential to the overall effectiveness and credibility of this methodology as a validation approach. Sufficient coverage should be with respect the ADS feature or ODD.

[1] A trip is a traversal of an entire travel pathway by a vehicle from the point of origin to a destination.

Coverage can be measured across different domains, and metrics can be used to determine sufficiency.

22. When validating the safety of an ADS, it is recommended that each scenario selected to test the ADS reflects the particular conditions (e.g., road configurations, direction of traffic in a given lane, etc.) that constitute the ODD in which the ADS is designed to operate. Scenarios should be relevant to the ADS feature being validated. For example, an ADS feature intended only for highway use would not be subject to a scenario involving turns at intersections.

23. Because an ADS will need to be responsive to actions by other road users, which may make a crash unavoidable, it is recommended that scenarios are not limited to those that are deemed preventable by the ADS. Unsafe behaviours of other road users (e.g. vehicle travelling in the wrong direction, sudden unsignalled lane changes, and exceeding the speed limit) —if reasonably foreseeable within the appropriate ODD—should be included as part of validation testing.

24. Consideration should be given to the many approaches that can be used to identify scenarios for safety validation purposes, including:

(a) Analysing human driver behaviour, including evaluating naturalistic driving data;

(b) Analysing collision data, such as law enforcement and insurance companies' crash databases;

(c) Analysing traffic patterns in specific ODD (e.g., by recording and analysing a road user behaviour at intersections);

(d) Analysing data collected from ADS' sensors (e.g., accelerometer, camera, radar, and global positioning systems);

(e) Using a specially configured measurement vehicle, onsite monitoring equipment, drone measurements, etc. for collecting various traffic data (including other road users);

(f) Knowledge/experience acquired during ADS development;

(g) Synthetically generated scenarios from key parameter variations;

(h) Engineered scenarios based on functional safety requirements and safety of intended functionality;

(i) composing complex scenario from existing catalogues of basic scenarios; and

(j) Random variations of all scenario parameters, both for the ADS an ORUs.

C. Classifying scenarios

25. The amount of information that is included in a scenario can be extensive. For example, the description of a scenario could contain information specifying a wide range of different actions, characteristics and elements[2], such as objects (e.g., vehicles, pedestrians), roadways, and environments, as well as pre-planned courses of action and major events that should occur during the scenario. Therefore, it is critical that a standardized and structured language for describing scenarios is established so that ADS stakeholders understand the intention of a scenario, each other's objectives, and the capabilities of an ADS. One tool for establishing uniform language for describing a scenario is a template, which ensures that the information to be included in the scenario is consistent and minimizes the possibility of confusion in its interpretation.

[2] Traffic scenarios are derived by combining a number of relevant elements describing the scenario space systematically.

26. It is recommended that a uniform language be used to describe a scenario to ensure that the information included is consistent and minimizes the possibility of confusion in its interpretation.

27. [It is recommended to describe scenarios by different levels of abstraction. Abstraction supplies the ability to focus the scenario description on specific aspects, while leaving other details for further processing as needed. Some Industries and researches are proposing 3 or 4 levels of scenario abstraction: Functional, Abstract, Logical, and Concrete. The essence of these levels is described below. The 3 or 4 levels do not imply nor mandate any specific implementation or translation flow from one level to the other.]

(a) Functional Scenario: [A scenario described in natural language on a conceptional level, in general without specific physical values. These are scenarios / Scenarios] with the highest level of abstraction, outlining the core concept of the scenario, such as a basic description of the ego vehicle's actions; the interactions of the ego vehicle with other road users and objects; and other elements that compose the scenario (e.g. environmental conditions etc.). This approach uses accessible language to describe the situation and its corresponding elements.

[(b) Abstract Scenario: A formalized, declarative description of the scenario, derived from functional scenario. The declarative specification on the abstract level enables highlighting of the relevant aspects of the scenario while focusing on efficient description of relations (Cause-effect).]

(c) Logical Scenario: [A scenario described with the inclusion of parameters, where the values of some of the parameters are defined as ranges.] For example, building off the elements identified within the functional scenario, developers generate a logical scenario by selecting value ranges or probability distributions for each element within a scenario (e.g., the possible width of a lane in meters).

(d) Concrete Scenarios: [A scenario depicted with explicit parameters values, describing physical attributes.] Concrete scenarios are established by selecting specific values for each element. This step ensures that a specific test scenario is reproducible. In addition, for each logical scenario with continuous ranges, any number of concrete scenarios can be developed, helping to ensure a vehicle is exposed to a wide variety of situations.

The following figures represents different options of using the levels of abstractions in order to derive concrete scenarios, other implementations are also possible.

[Figure 1-1
Examples of a scenario using functional, logical and concrete categorizations (Pegasus, 2018)

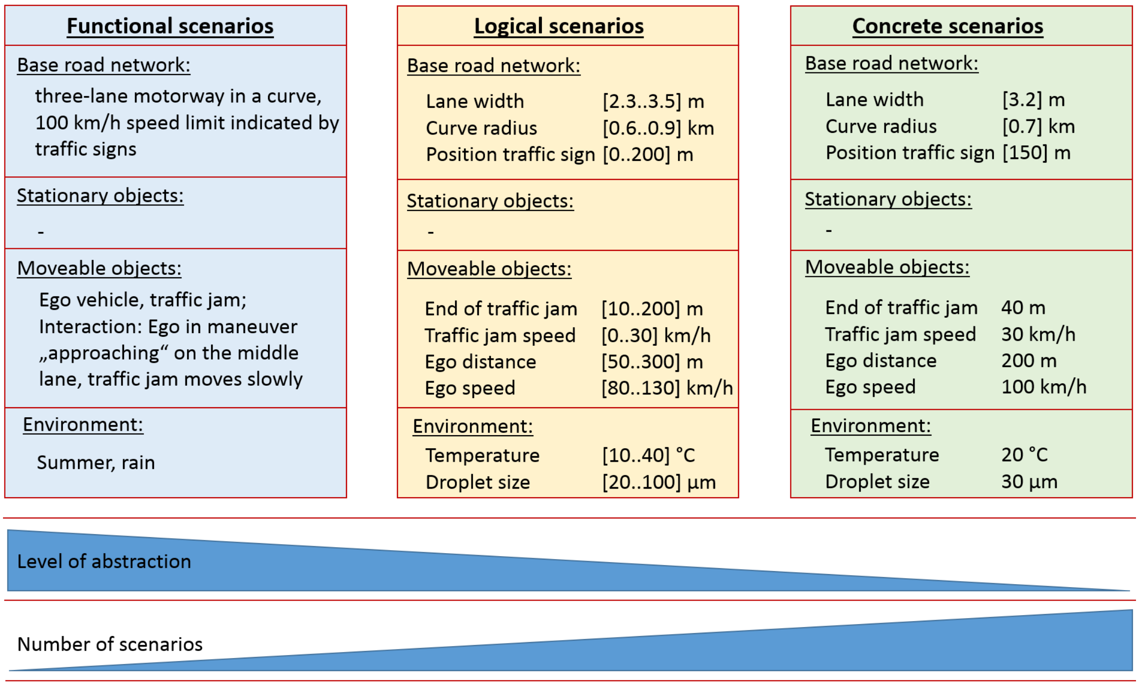
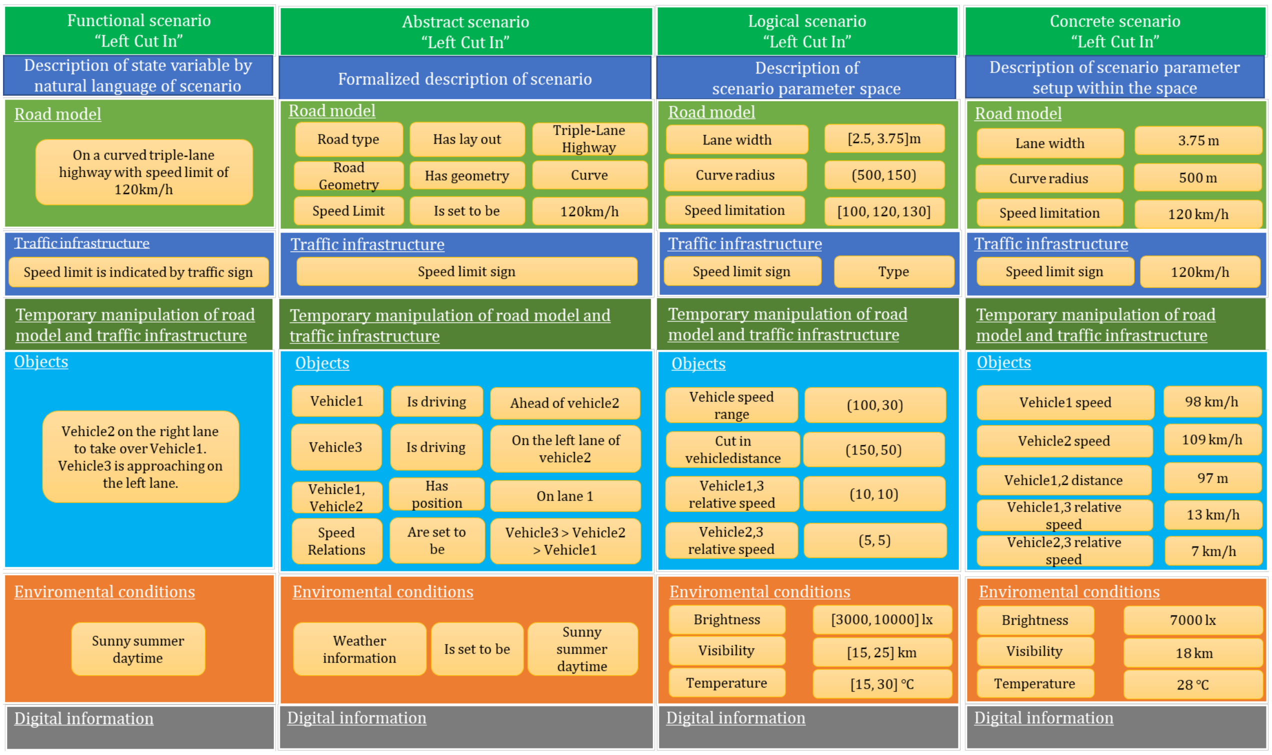


Figure 1-2

Examples of the relationship of functional scenario, abstract scenario, logical scenario and concrete scenario (ISO 34501)]



28. To provide some illustrative examples, VMAD has prepared a series of functional scenarios for a divided highway application, which are described in Annex II. As previously mentioned, this guidance document should be regarded as an "evergreen document". As such, this annex will be updated based on ongoing discussions by VMAD and other WP.29 working groups. It is anticipated that future iterations of Annex II will also incorporate scenarios with lower levels of abstraction (e.g. logical scenarios and potential approaches for describing them). As previously noted, VMAD will also continue to examine the development of a more comprehensive scenarios catalogue as part of the NATM.

D. Scenario usage

29. The use of scenarios can be applied to different testing methodologies, such as virtual/simulation, test track, and real-world testing. Together, these methodologies provide a multifaceted testing architecture, with each methodology possessing specific strengths and weaknesses. Therefore, some scenarios may be more appropriately tested using certain test methodologies over others.

30. It is recommended that sampling techniques be used when selecting parameters to be used in creating logical and concrete scenarios for ADS validation for a particular ADS and its ODD to avoid the ADS being optimized for a set of known test cases. Using a maxim number of random samples is clearly preferable from a credibility perspective, it is recognized that this can place a greater burden on manufacturers and the relevant authority (e.g. technical service). This should be considered when determining the volume of tests to be conducted when using the random sampling. It is assumed that for simulation/virtual testing the burden of random sampling is less and therefore maximizing the number of random samples for this facet of the testing is more feasible.

VI. Simulation/virtual testing – Pillar 1

A. Types of simulation toolchain approaches

31. The simulation toolchain used for virtual testing may result in the combination of different approaches. In particular, there are many ways that tests can be performed:

(a) Entirely inside a computer (referred to as Model or Software in the Loop testing, MIL/SIL), with the model of the elements involved (e.g., a simple representation of the control logic of an ADS) interacting in a simulated environment; and/or

(b) With a sensor, a subsystem, or the whole vehicle interacting with a virtual environment (Hardware or Vehicle in the Loop testing, HIL/VIL). For VIL testing, the vehicle can either be in:

(i) A laboratory where the vehicle would be standing still or moving on a chassis dynamometer or on a powertrain test bed and is connected to the environment model by wire or by direct stimulation of its sensors; or

(ii) A *proving ground* where the vehicle would be connected to an environment model and would interact with virtual objects by physically moving on the *test-track*.

(c) With a subsystem interacting with a real driver (*Driver in the Loop testing*, *DIL*).

32. Interaction between the system and the environment; The interaction between the system under the test and the environment can either be an open- or closed-loop.

33. An open-loop test (also referred to as software or hardware reprocessing, shadow mode, etc.) could be done through a variety of methods, such as the ADS interacting with recorded real world data. In this case, the virtual objects' actions are data-driven only and the

information is not self-corrected based on feedback from the output. Because the open-loop controller actions may vary due to external disturbances without the ADS and/or the assessor being aware, the applicability of open-loop tests in the ADS validation may be limited.

34. Closed-loop virtual tests include a feedback loop that continuously sends information from the "closed-loop" controller back to the ADS when the ADS takes an action. Within these test systems, the digital objects in the environment could react in different ways depending on the action of the system under test.

35. Selecting an open- or closed-loop test could depend on factors such as the objectives of the virtual testing activity and the status of development of the system under test. For ADS validation it is expected that mainly closed-loop virtual testing will be considered.

36. The flexibility of simulation makes it a standard test method during a vehicle's design and the development of this pillar will also make it part of the ADS validation process. For an ADS, it will be impossible to test the vehicle's behaviour in the real world for all possible situations as well as for any subsequent change in the ADS' driving logic. Virtual testing will therefore become an indispensable tool to verify the capability of the automated system to deal with a wide variety of possible scenarios. In addition, virtual testing can be beneficial in replacing real world and proving ground testing where there are concerns over safety-critical traffic scenarios. It is recommended therefore that virtual testing be used to test the ADS under safety critical scenarios that would be difficult and/or unsafe to reproduce on test tracks or public roads.

37. Virtual tests used for ADS validation can achieve different objectives, depending on the overall validation strategy and the accuracy of the underlying simulation and models.

(a) Provide qualitative confidence in the safety of the full system.

(b) Contribute directly to statistical confidence in the safety of the full system (caveats apply).

(c) Provide qualitative or statistical confidence in the performance of specific subsystems or components.

(d) Discover challenging scenarios that can be tested in the real world (e.g. real-world tests and track tests described in chapter 7 and 8 of this document).

38. In contrast to all its potential benefits, a limitation, of this approach, is in its intrinsic limited fidelity. As models provide a representation of the reality, the suitability of a model to satisfactorily replace the real world for validating the safety of an ADS has to be carefully assessed. Therefore, the validation of the simulation and models used in virtual testing is essential to determine the quality and reliability of the results compared to real-world performance.

39. It is recommended that a virtual test of the ADS' performance is compared with its performance in the real world when executing the same scenario. This will provide the opportunity to assess the accuracy of the virtual testing toolchain that is used. Given the high number of scenarios that virtual testing can perform compared to track testing, the validation will probably need to be performed on a smaller but still sufficiently representative subset of the relevant scenarios in order to substantiate any extrapolation beyond the scenarios used for the validation. More information as well as recommendations on credibility assessment for using virtual toolchains can be found in Annex III.

40. In the short-term virtual testing may only be conducted using simulation toolchains developed and maintained by the ADS manufacturer. Since their design depends on the validation and verification strategies implemented by the manufacturer, it is recommended that simulation toolchains are not subject to regulation or standardization at this time. Rather, simulation toolchains should be explained and documented by the ADS manufacturer and its suitability assessed during the certification process. For this reason, the output of the NATM related to virtual testing ensures that documentation and data provided by the manufacturer is appropriate. Furthermore, virtual testing using modelling and simulation should be credible enough for an assessor to make sound decisions. Credibility is discussed further below.

41. It is recommended that when validating the safety of the ADS, particular attention should be placed on the interaction between virtual testing and the other test methods. Virtual testing will have strong relationships with all the pillars of the NATM guidelines. In particular:

(a) Virtual testing supplements physical testing to account for the quantity and diversity of ADS configurations, intended uses and limitations on use. One of the strengths of virtual testing is its capacity to assess the ADS performance across multiple scenarios and across ranges of parameters within scenarios in a cost-effective manner. Virtual testing enables results of limited physical tests to be supplemented by verifiable data covering numerous instances of the test scenario, by varying parameters. Using this approach, virtual testing can demonstrate ADS coverage of safety-critical scenarios, and hence provide evidence that an ADS will perform as intended for that type of scenario in the real world. These advantages reduce the burden on physical tests (offsetting their weaknesses) and help to improve the efficiency of the overall assessment process across the pillars. Virtual testing can also be effectively used to identify and cover edge cases and other low-probability scenarios to increase confidence on the ADS' likely performances.

(b) Virtual testing can play an important role in the development of traffic scenarios.

(c) Virtual testing enables assessment of ADS performance boundaries, enabling precise definition of the boundaries between collision avoidance and crash mitigation. Through methods of randomization and scenario compositions, virtual testing enables the developer or the assessor to challenge the ADS and increase confidence in its performance when challenged with low probability events.

(d) Virtual testing will be a key element in the audit assessment. Results of virtual testing carried out both during vehicle development and in the verification and validation phase will provide valuable evidence supporting the safety audit. The manufacturers will need to provide evidence and documentation about how the virtual testing is carried out and how the underlying simulation toolchain has been validated.

(e) Results from real-world tests can improve the accuracy of simulation and models.

(f) Virtual testing can play an important role in responding to concerns identified through in-use monitoring of ADS performance. Virtual testing provides a quick and flexible approach to analyse ADS performance based on real-world events. It allows manufacturers to understand and verify the ADS behaviour and to understand why an issue may have occurred. It may identify an untested scenario, or a set of untried parameters. It may also identify the "scale" of any issue. If the virtual testing does identify unsafe behaviour it can then also help to assess the efficacy of modifications to the ADS and ultimately to improve the overall ADS performance. Where appropriate, the information and scenario descriptions can be shared and integrated into scenarios and testing regimes worldwide.

42. It is recognised that specific regulatory functional safety requirements are still under development. Virtual testing however, using a validated simulation toolchain, shows promise for assessing the following general safety requirements that are currently under consideration:

(a) The ADS should drive safely and manage safety critical situations. These are the requirements where virtual testing can play a prominent role. MIL/SIL, HIL and VIL virtual testing can all be used to assess these requirements at different stages of vehicle verification and validation.

(b) The ADS should interact safely with the user. DIL virtual testing can be helpful to support the assessment of this category of safety requirement by analysing the interaction between the driver and the ADS in a safe and controlled environment.

(c) The ADS should safely manage failure modes and ADS should ensure a safe operational state. The use of virtual testing in these two categories is also very promising but would probably require further research work. SIL virtual testing could include simulated failures and maintenance requests. HIL and VIL virtual testing could be used to assess how the system would react to the occurrence of a malfunctioning induced into the real system.

VII. Track testing – Pillar 2

43. Track testing occurs on a closed-access testing ground that uses real obstacles and obstacle surrogates (e.g., vehicle crash targets, etc.) to assess the safety requirements of an ADS (e.g., human factors, safety system). This testing approach allows for the physical vehicles to be tested through realistic scenarios to evaluate either sub-systems or the fully assembled system. These external inputs and conditions can be controlled or measured during a test.

44. Track testing is suitable for assessing the ADS capabilities in nominal scenarios and critical scenarios. The same tests can be used to verify the performance of the vehicles regarding human factors or fallback in these scenarios. However, operating on test tracks can be resource intensive. For more background information on track testing, such as its strengths and weaknesses, please review the NATM Master Document.

45. It is recommended that track testing be used to assess the performance of ADS in a number of selected important nominal and critical scenarios, notably given that, unlike real-world testing, track testing can accelerate exposure to known rare events or safety critical scenarios, and in a more controlled and safer environment.

46. It is furthermore recommended to develop the track tests in line with the approach set out in Annex V.

A. How the pillar interacts with other pillars

47. It is recommended that information generated during the track-test be used as additional data to validate the virtual tests by comparing an ADS' performance between a virtual test and a test track on the same scenario. For instance, track testing can be used as an additional tool/method to validate the quality/reliability of the virtual toolchain. However, it is important to keep in mind the limitations described in the NATM Master Document.

VIII. Real-world testing – Pillar 3

48. Real-world testing uses public roads to test the capabilities and compliance with safety requirements (e.g., human factors, safety system) of a vehicle with an automated driving system (ADS) in real-world traffic. It therefore provides an opportunity to validate the safety of the ADS within its true operating environment. For more background information on real world testing, such as its strengths and weaknesses, please review the NATM Master Document.

49. It is recommended that real world testing:

(a) be considered for assessing aspects of the ADS performance related to its capability to drive in real traffic conditions, e.g. smooth driving, capability to deal with dense traffic, interaction with other road users, maintaining flow of traffic, being considerate and courteous to other vehicles;

(b) be considered for assessing aspects of the ADS performance at some ODD boundaries (nominal and complex scenarios), i.e. is the system triggering transition demands to the driver when it is supposed to (e.g. end of the ODD, weather conditions). The same testing could be used to confirm the performances related to human factors under these conditions;

(c) be considered for detecting issues that may not be well captured by track tests and simulation, such as perception quality limitation (e.g. due to light conditions, rain, etc.): and,

(d) be developed in line with the approach set out in Annex V.

53. Although it may not be possible to encounter all traffic scenarios during a real-world test, the likelihood of covering specific complex scenarios could be increased by selecting a specific type of ODD (e.g., highway) and examining when and where specific elements (e.g., high- or low-density traffic) typically occur.

54. Specific infractions identified during real-world testing may be reviewed and/or assessed by evaluating the data gathered during that test and any data gathered during additional virtual, track and real-world testing.

A. How the pillar interacts with other pillars

55. Data generated during real-world testing may be used as additional data to validate whether portions of a virtual and/or track-testing environment were modelled properly by comparing an ADS' performance within a simulation and/or track test with its performance in a real-world environment when executing the same test scenario.

56. It can also be used to support the development of new traffic scenarios for track and virtual testing, allowing for the identification of edge cases and other unanticipated hazardous situations that could challenge the ADS.

57. The information gathered from real world testing may also support improvements in the hazard and risk analysis and design of the ADS systems.

IX. Audit – Pillar 4

58. The purpose of the audit pillar is to assess/demonstrate that the:

(a) Manufacturer has the right processes to ensure operational and functional safety during the vehicle lifecycle, and

(b) Vehicle's design is safe by design and that the design has been sufficiently validated before market introduction.

58 bis. Therefore, this pillar is composed of two main components: one is the audit of the manufacturer processes established through a safety management system, and the other consists of the safety assessment of the ADS design.

59. It is recommended that the manufacturer is required to demonstrate that:

(a) Robust processes are in place to ensure safety throughout the vehicle's lifecycle (development, production, operation and decommissioning). This shall include taking the right measures to monitor the vehicle during the in-service operation and to take appropriate (corrective or preventive) action to address any issues;

(b) The hazards and risks of the ADS have been identified and it is clear that a "safety-by-design" approach exists and had been applied to mitigate them; and

(c) The risk assessment and the safety-by-design approach have been validated, through testing, by the manufacturer and show that the vehicle meets the safety requirements before market introduction. The vehicle should be free of unreasonable safety risks to the broader transport ecosystem, and in particular, to the driver, passengers and other road users.

60. Based on the evidence provided by the manufacturer and including the tests, authorities will be able to assess whether the processes, the risk assessment, the design and the validation are robust enough with regard functional and operational safety.

A. General guidance on the audit of the manufacturer safety management system

61. The purpose of the audit of the manufacturer's safety management system is to confirm that the manufacturer has robust processes to manage safety risks and to ensure safety throughout the ADS lifecycle (development, production, operation and

decommissioning). It should include taking appropriate measures to monitor the vehicle during the in-service operation and to take the corrective remedial action when necessary.

62. The documentation provided by a manufacturer should demonstrate that their safety management system has effective processes, methodologies and tools. It should be up to date and also clear that it is being used within the organization. It should show how the organization intends to manage safety and to demonstrate continued compliance throughout the product lifecycle (design, development, production, operation and decommissioning).

1. Safety Management System (SMS)

63. An SMS is a systematic approach to managing safety, which encompasses and integrates organizational, human and technical factors:

(a) Human component ensuring the ADS lifecycle leveraged upon personnel with appropriate skills, training, and understanding to identify risks and appropriate mitigation measures,

(b) Organisational component procedures and methods that help to manage the identified risks, understand their relationships and interactions with other risks and mitigation measures. Helping to ensure that there are no unforeseen consequences.

(c) Technical component using appropriate tools and equipment.

63 bis. An adequate SMS will incorporate, monitor and improve all three dimensions and help to control the identified risks. The SMS evaluation is based on automotive engineering standards, guidebooks, and best practice documents relevant to safety.

64 It is recommended that any operational risk identified in the product should, where appropriate, have mitigations implemented during the Design and Development phase. The ADS manufacturer should then be able to show the link between the overall risk management process, the mitigations and the resulting operational risks.

65. Examples of processes and activities that are recommended to be documented by the manufacturer:

(a) Risk Management:

(i) Risk identification (in line with ISO 31000 para. 6.4.2 standard or equivalent)

(ii) Risk analysis (in line with ISO 31000 para. 6.4.3 standard or equivalent)

(iii) Risk evaluation (in line with ISO 31000 para. 6.4.4 standard or equivalent)

(iv) Risk treatment (in line with ISO 31000 para. 6.4.5 standard or equivalent),

(v) Processes for keeping the risk assessments up to date,

(vi) Safety performance of the organization and effectiveness of safety risk controls.

(b) Safety Governance

(i) Safety policies and principles (in line with the concept stated in ISO 21434, para. 5.4.1 and ISO 9001 Automotive 5.2, but from a safety perspective)

(ii) Management commitment (in line with the concept stated in ISO 21434, para. 5.4.1 and ISO 9001 Automotive 5.1, but from safety perspective)

(iii) Roles and responsibilities (ISO 26262-2, para. 6.4.2, this relates to the organizational and project dependent activities)

(c) Safety culture (ISO 26262-2, para. 5.4.2)

(d) Effective communications within the organization (ISO 26262-2, para. 5.4.2.3)

(e) Information sharing outside of the organization (in line with the concept stated in ISO 21434, para. 5.4.5 and ISO 9001, but from a safety perspective)

(f) Quality Management System (e.g., as per IATF 16949 or ISO 9001 or equivalent) to support safety engineering, including change management, configuration management, requirement management, tool management etc.

66. It is recommended that the design and development process is well established and documented. It should include risk management, requirements management, requirements' implementation, testing, failure tracking, remedial actions, and release management.

67. Examples of processes and activities that should be documented to ensure the robustness of the design and development phase:

(a) A general description of how the organization performs all the design and development activities

(b) Vehicle\system development, integration, and implementation.

(i) Requirements management (e.g. Requirement capture and validation)

(ii) Validation strategies, including but not limited to

a. Assessment of the physical testing environment

b. Credibility assessment for virtual tool chain

c. System integration

d. Software

e. Hardware

(iii) Management of functional Safety and operational safety, including the ongoing evaluation and update of risk assessments and interactions with In-Service Safety

(iv) Management of Human Factors (e.g. Human centered design processes)

(c) Design and change management, including but not limited;

(i) The major design decisions,

(ii) The relevant design modifications to the ADS

(iii) The personnel involved in the design

(iv) The tools and thresholds adopted for the ADS safety verification.

68. It is recommended that the manufacturer institutes and maintains effective communication channels between the departments responsible for functional/operational safety, cybersecurity and any other relevant disciplines related to the achievement of vehicle safety.

69. Examples of processes and activities that should be considered to assure that responsibilities are properly discharged:

(a) Roles and responsibilities of the people involved during the design and development phase

(b) Qualifications and experience of persons responsible for making decisions that affect safety

(c) Coordination of roles, responsibilities and information transfer between design and production activities

70. Examples of processes and activities that are recommended to be documented to ensure the robustness of the development and the production phase include:

(a) Quality Management System accreditation (e.g., as per IATF 16949 or ISO 9001 or equivalent)

(b) A description of the way in which the organisation performs all the production functions including management of working conditions, working environment, equipment and tools.

71. Examples of processes and activities to be documented to assure robustness of development and distributed production:

(a) Liaison between the vehicle and/or ADS manufacturer and all other organisations (partners or subcontractors) involved

(b) Criteria for the acceptability of “subsystem/components” manufactured by other partners or subcontractors. (i.e., deployment of production assurance requirements to supply chain)

72. It is recommended that the manufacturer demonstrate that periodic independent internal audits and external audit are carried out to ensure that the processes established for the Safety Management System are implemented consistently. (UN R157, para. 3.5.5, ISO 26262-2, para. 6.4.11)

73. The following are examples of processes and activities that should be documented to assure independent design audit and assessment:

(a) assurance that all practices and procedures applied during the vehicle\system development are followed;

(b) assurance that there is an independent check of compliance with the applicable requirements and regulations is performed. (i.e., not from person creating the compliance data);

(c) process to assure the continuing evaluation of the Safety Management System to ensure that it remains effective.

74. It is recommended that a manufacturer puts in place suitable arrangements (e.g. contractual arrangements, clear interfaces, quality management system) with any organization involved in the development, manufacturing or in-use deployment of their vehicles (e.g. contracted suppliers, service providers or manufacturers’ sub-organizations) to ensure that their approach to safety management related to the committed activities complies with the recommendations of the present guidelines.

75. Examples of processes and activities that are recommended to be documented:

(a) Organizational policy for supply chain

(b) Incorporation of risks originating from supply chain

(c) Evaluation of supplier SMS capability and corresponding audits

(d) Processes to establish contracts, agreements for ensuring safety across the phases of development, production, and postproduction

(e) Processes for distributed safety activities.

2. Link with the in-service monitoring/reporting pillar

76. It is recommended that a manufacturer has processes to monitor safety-relevant incidents/ crashes/collisions caused by the ADS. The manufactures should also have a process to manage potential safety-relevant gaps during the in-service operation phase (possibly identified by in-service monitoring) and a process to update those vehicles.

77. The manufacturer should have processes to report safety relevant occurrences (e.g. collision with another road users and potential safety-relevant gaps, see the In-service Monitoring and Reporting Pillar) to the relevant authority when they occur.

78. The manufacturers should set up processes for the operational phase to confirm of compliance with the defined safety case. It should include, early detection of new unknown situations (in line with SOTIF safety development goal to minimize the unknown scenarios area), event investigation, to share learnings derived from incidents and near-miss analysis

to allow the whole community to learn from operational feedback and to contribute to the continuous improvement of automotive safety

79. Example of guiding principles: Is there a document describing the appropriate procedure of reporting incidents to the management? Is there evidence that the company is complying with that procedure? Is there a document describing the appropriate procedure of investigation and documentation of incidents? Is there evidence that the company is complying with that procedure?

3. Updates of the SMS

80. SMS documentation shall be regularly updated in line with any relevant changes to the SMS processes. It is recommended that gap analysis should be used when auditing and updating the SMS, examining the current safety culture before formulating new and more appropriate SMS processes to ensure issues are adequately resolved. The SMS shall be subject to a process of continual improvement (e.g. "Plan, Do, Check, Act as described in ISO 9001). Any changes to SMS documentation should be communicated as required to the relevant authority.

B. General guidance on the safety assessment of the ADS design

81. The purpose of the audit of the safety by design concept of the ADS is to demonstrate that hazards and risks relevant to the ADS have been identified by the manufacturer and a consistent safety-by-design concept has been implemented to mitigate these risks. In addition, it should demonstrate that the risk assessment and the design have been validated by the manufacturer through testing. This should demonstrate that, before the vehicle is placed on the market, it meets the relevant safety requirements. This means it is free of unreasonable safety risks to the broader transport ecosystem and in particular to the driver, passengers and other road users.

1. ADS General Description

82. It is recommended that a description should be provided, which gives a simple explanation of the operational characteristics of the ADS and ADS features:

(a) Operational Design Domain (Speed, road type, country, Environment, Road conditions, etc.);

(b) Basic performance (e.g. Object and Event Detection and Response (OEDR), etc.)

(c) Interaction with other road users

(d) Main conditions for Minimum Risk Manoeuvres.

(e) Interaction with the driver (if relevant)

(f) Supervision centre (if relevant)

(g) The method of activating, overriding or deactivating the ADS by any or all of the driver (where relevant), the human supervision centre (where relevant), passengers (where relevant) or other road users (where relevant).

2. Description of the functions of the ADS

83. A description should be provided which gives a clear explanation of all the functions including control strategies of the ADS and the methods employed to perform the dynamic driving tasks within the ODD and the boundaries under which the ADS is designed to operate, including a statement of the mechanism(s) by which control is exercised.

84. It is recommended that a list of all input and sensed variables is provided and the working range of these defined, along with a description of how each variable affects system behaviour.

85. A list of all output variables which are controlled by the ADS should be provided and an explanation given, in each case, of whether the control is direct or via another vehicle system. The range of control exercised on each variable should be defined.

3. ADS layout and schematics

(a) Inventory of components

86. A list should be provided, including all the units of the ADS and mentioning the other vehicle systems which are needed to achieve the control function in question.

87. An outline schematic showing these units and their relationships should be provided, with both the equipment distribution and the interconnections made clear.

88. It is recommended that the outline includes:

(a) Perception and objects detection including mapping and positioning

(b) Characterization of decision-making

(c) Remote supervision and remote monitoring by a remote supervision centre (if applicable).

(d) Information display / user interface

(e) The data storage system (e.g., DSSAD).

(b) Functions of the units

89. The function of each unit of the ADS should be outlined and the signals linking it with other units or with other vehicle systems should be shown. This may be provided by a labelled block diagram or other schematic, or by a description aided by such a diagram.

90. It is recommended that interconnections within the ADS should be shown by a circuit diagram for the electric transmission links, by a piping diagram for pneumatic or hydraulic transmission equipment and by a simplified diagrammatic layout for mechanical linkages. The transmission links both to and from other systems should also be shown.

91. There should be a clear correspondence between transmission links and the signals carried between units. Priorities of signals on multiplexed data paths should be stated wherever priority may be an issue affecting performance or safety.

(c) Identification of units

92. Each unit should be clearly and unambiguously identifiable (e.g. by marking for hardware, and by marking or software identification for software content). This should provide a clear method for identifying the hardware and software in the associated documentation. Where the software version can be changed without requiring replacement of the marking or component, the software identification must be updated by means of the newly released software.

93. It is recommended that where functions are combined within a single control unit or indeed within a single computer, but shown in multiple blocks in the diagram, then for clarity and ease of explanation, only a single hardware identification marking should be used.

94. The identification defines the hardware and software version and, where the software changes and alters the function of the unit, the identifier associated with that software should also be changed.

(d) Installation of sensing system components

95. The manufacturer should provide information regarding the installation options that will be employed for the individual components that comprise the sensing system. These options should include, but are not limited to, the location of the component in/on the vehicle, the material(s) surrounding the component, the dimensioning and geometry of the material surrounding the component, and the surface finish of the materials surrounding the component, once installed in the vehicle. The information should also include installation

specifications that are critical to the ADS's performance, e.g., tolerances on installation angle.

96. Any changes to the individual components of the sensing system, or the installation options, should be updated in the documentation.

(e) ADS specifications

(a) Description of ADS specifications in Normal and Emergency Conditions, acceptance criteria and the demonstration of compliance with those criteria.

(b) List of applied regulations, codes, and standards

(f) Safety Concept and validation of the safety concept by the manufacturer

97. The manufacturer should provide a statement which affirms that the ADS is free from unreasonable risks for the driver (if applicable), passengers and other road users.

98. In respect of software employed in the ADS, the outline architecture should be explained and the design methods and tools used should be identified. The manufacturer should show evidence of how the ADS capabilities were realized and checked during the design and development process.

99. It is recommended that the manufacturer should provide an explanation of the design provisions built into the ADS to ensure functional and operational safety. Possible design provisions in the ADS include:

(a) Fall-back (or fail safe) operation using a partial system.

(b) Redundancy using separate systems.

(c) Removal of some or all automated driving function(s).

100. If a chosen provision utilizes a partial performance mode of operation under certain fault conditions (e.g. in case of severe failures), then these conditions should be stated (e.g. type of failure). The resulting ADS behaviour and capabilities should be defined (e.g. initiation of a minimum risk manoeuvre immediately) as well as the warning strategy to the driver/remote supervision centre (if applicable).

101. If the chosen provision selects a second (back-up) means to realize the performance of the dynamic driving task, it is recommended that the principles of the change-over mechanism, the logic and level of redundancy and any built-in back-up checking features be explained and the resulting limits of back-up effectiveness defined.

102. If the chosen provision selects the removal of an automated driving function, it is recommended that this is done in compliance with the relevant provisions of this regulation. All the corresponding output control signals associated with this function should be inhibited.

103. The documentation should be supported, by an analysis which shows how the ADS will behave to mitigate or avoid hazards which can have a bearing on the safety of the driver (if applicable), passengers and other road users. It should show how unknown hazardous scenarios will be managed by the manufacturer to keep the residual risk level under control.

104. The chosen analytical approach(es) should be established by the manufacturer and made available for assessment to the relevant authority before market introduction. The auditor should perform an assessment of the application of these analytical approach(es), including:

(a) Inspection of the safety approach at the concept (vehicle) level.

(b) It is recommended that this approach be based on a Hazard / Risk analysis appropriate to system safety.

(c) Inspection of the safety approach at the ADS level including a top down (from possible hazard to design) and bottom-up approach (from design to possible hazards). The safety approach may be based on a Failure Mode and Effect Analysis (FMEA), a Fault Tree Analysis (FTA) and a System-Theoretic Process Analysis (STPA) or any similar process appropriate to system functional and operational safety.

(d) Inspection of the documentation that should demonstrate the validation/verification plans and results including appropriate acceptance criteria. It should include testing appropriate for validation, for example, Hardware in the Loop (HIL) testing, vehicle on-road operational testing, testing with real end users, or any other testing appropriate for validation/verification.

105. The auditor/assessor should perform an assessment of the physical testing (proving ground and/or public road) environment and should assess the documentation of the virtual tool chain provided by the manufacturer. The auditor/assessor may decide to carry out tests of the complete integrated tool to assess the credibility of the virtual tool chain.

106. Results of validation and verification may be assessed by analysing coverage of the different tests and setting minimal coverage thresholds for various metrics.

107. It is recommended that the documentation confirms that at least each of the following items are covered where applicable:

(a) Issues linked to interactions with other vehicle systems (e.g., braking, steering);

(b) Failures of the automated driving system and the resulting risk mitigation strategy;

(c) Situations within the ODD when a system may create unreasonable safety risks for the driver (if applicable), passengers and other road users due to operational disturbances, for instance:

- lack of or wrong comprehension of the vehicle environment;
- lack of understanding of the reaction from the driver (if applicable), passenger or other road users;
- inadequate control;
- challenging scenarios.

(d) Identification of the relevant scenarios within the ODD boundaries and the methodology used to select scenarios and choose the validation methodology and approach.

(e) Decision making process for the performance of the dynamic driving tasks (e.g. emergency manoeuvres), the interaction with other road users and the compliance with traffic rules

(f) Cyber-attacks that may have an impact on the safety of the vehicle.

(g) Reasonably foreseeable misuse by the driver (if applicable) (e.g., the use of a driver availability recognition system and an explanation on how the availability criteria were established), mistakes or misunderstanding by the driver if applicable (e.g., unintentional override) and intentional tampering of the ADS.

108. The documentation should have arguments supporting the safety concept that is understandable and logical and cover all the different functions of the ADS.

109. The documentation should also demonstrate that validation plans are robust enough to demonstrate safety (e.g., reasonable coverage of chosen scenarios as part of the validation methodology chosen) and have been completed.

110. It is recommended that the documentation provides evidence that the vehicle is free from unreasonable risks for the driver (if applicable); vehicle occupants and other road users in the operational design domain. This could be achieved through:

(a) An overall validation target (i.e., validation acceptance criteria) supported by validation results, demonstrating that at entry into service of the ADS will not increase the overall level of risk for the driver (if applicable), vehicle occupants, and other road users compared to a manually driven vehicles; and

(b) A scenario specific approach showing that the ADS will not increase the overall level of risk for the driver (if applicable), passengers and other road users compared to a manually driven vehicles for each of the safety relevant scenarios.

111. The documentation should allow the relevant authority to test and verify the safety concept.

112. It is recommended that the documentation itemizes the parameters being monitored on the vehicle and should set out, for each failure condition of the type defined in accordance with 84.6. of this annex, the warning signal to be given to the driver (if applicable) /vehicle occupants/other road users and/or to service/technical inspection personnel.

113. This documentation should also describe the measures in place to ensure the ADS is free from unreasonable risks for the driver (if applicable), vehicle occupants, and other road users when the performance of the ADS is affected by environmental conditions e.g. climatic, temperature, dust ingress, water ingress, ice packing.

(g) Data Storage System

114. It is recommended that the documentation describe:

(a) Storage location and crash survivability

(b) Data recorded during vehicle operation and occurrences

(c) Data security and protection against unauthorized access or use

(d) Means and tools to carry out authorized access to data.

(h) Cyber Security & Software Update Management

115. The documentation should describe:

(a) Cyber security and software update management,

(b) Identification of risks, mitigation measures,

(c) Secondary risks and assessment of residual risks,

(d) Software update procedure and management put in place to comply with legislative requirements.

(i) Information provisions to users

116. It is recommended that the documentation includes:

(a) The distinction between maintenance and an operational manual,

(b) A safety precaution manual that includes safety-relevant information for the user,

(c) A briefing on the user's role and how it might change during the vehicle operation, including when the user is responsible for the safety and control of the vehicle,

(d) Information on how to use the ADS,

- o Transition of Control (ToC), where applicable
- o Take over
- o ADS activation
- o ODD
- o Role of the user after regaining control

(e) System Description and functional limitations

(f) Operational description (e.g., implications of switching off the ADS)

(g) Nominal Operations

(h) Emergency Operations

(i) Role of the user within the ADS' ODD

(j) Information related to the HMI's indications

- o Visual tell-tales, icons
- o Auditory signs
- o Haptic signs

(k) Means to deactivate the automated driving mode (take-over)

(l) Safety measures to be taken in the event of malfunctioning of the ADS

(m) Extent, timing and frequency of maintenance operations

(n) Means to enable a periodical technical inspection

(o) Documents and templates for maintenance, repair and periodical technical inspection

(p) Precautionary statements in the sense of compliance with limit values for the technical functions

(q) Data protection and data security functionalities

(r) List of system fault codes

(j) Safety management system

117. The manufacturer should have a valid Safety Management System relevant to the specific ADS and should inform the authority of any change that will affect the Safety Management System for the specific ADS.

(k) Type of documentation to be provided

118. Type of documentation to be provided by the manufacturer. Documentation should be brief yet provide evidence that the design and development has had the benefit of expertise from all the ADS fields which are involved.

(a) A documentation package which gives access to the basic design of the ADS and how it is linked to other vehicle systems or by which it directly controls output variables.

(b) Documentation explaining the function(s) of the ADS, including the control strategies and the safety concept.

(c) For periodic technical inspections, the documentation should describe how the current operational status of the ADS can be checked

(d) Documentation about how the software version(s) and the failure warning signal status can be readable in a standardized way via the use of an electronic communication interface (i.e., using a standard interface, such as the OBD port).

119. It is recommended that the documentation package shows that the ADS:

(a) Is designed and was developed to operate in such a way that it is free from unreasonable risks for the driver (if applicable), passengers and other road users within the declared ODD;

(b) Is capable of recognizing its boundaries;

(c) Respects any performance requirements specified by FRAV;

(d) Was developed according to the development process/method declared by the manufacturer;

120. Documentation should be made available in three parts:

(a) An information document which is submitted to the authority and should contain brief information on all the items.

(b) The formal documentation package annexed to the information document, which should be supplied to the Authority for the purpose of conducting the safety assessment.

(c) Additional confidential material and analysis data (intellectual property) which should be retained by the manufacturer, but made open for inspection (e.g. on-site in the engineering facilities of the manufacturer) at the time of the product assessment / process audit. The manufacturer should ensure that this material and analysis data remains available for a period of 10 years counted from the time when production of the ADS is discontinued.

121. Any changes to ADS safety design should be communicated as required to the relevant authority.

X. In-service monitoring and reporting – Pillar 5

122. The In-Service Monitoring and Reporting pillar (ISMR) addresses the in-service safety of automated vehicles after market introduction. In practice, the application of the other pillars of the NATM guidelines will assess whether the ADS is safe, according to the existing criteria, for market introduction; whereas the in-service monitoring and reporting will gather additional evidence from its in-service operation to demonstrate that the ADS continues to be safe after market introduction, i.e., that use of the ADS does not present an unreasonable safety risk. This pillar describes how to monitor the dynamic nature of the in-service operational use and then to provide feedback to ensure that there is continuous improvement of the safety of the ADS.

123. The pillar relies on the collection of relevant data during ADS operation.

124. This pillar does not address the obligation for "real-time monitoring" (self-checks/ on board diagnostics) of the performance of ADS subsystems by the manufacturer, which is part of the overall safety requirements. However, some form of monitoring of the performance of ADS subsystems over time could be part of "Objective 1" that is described in "General guidance on ISMR implementation" below, and could contribute to the predictive monitoring of safety performance degradation.

125. The processes put in place by the manufacturer to manage safety of the ADS during in-service operation, e.g. to manage changes in the traffic rules and in the infrastructure, fall outside this pillar and are assessed with the audit pillar. This pillar focuses on the type of data to be monitored and reported.

126. Whatever safety evaluation is done before market introduction, the actual level of safety will only be confirmed once enough vehicles have reached the in-service operation phase and have encountered a sufficient range of traffic and environmental conditions. It is recommended that a feedback loop (fleet monitoring) is put in place to confirm the safety argument and confirm the validation carried out by the manufacturer before market introduction. The operational experience feedback from in-service monitoring will allow ex-post evaluation of regulatory requirements and validation methods and could provide indications on safety related issues that need reviewing.

127. New safety risks might be identified and/or introduced during the in-service operation of ADS vehicles. The In-Service Monitoring and Reporting pillar can be used to identify them and provide data to update the common scenario catalogue to cover them.

128. Finally, in the early phase of market introduction of ADS, it is essential that the whole community can learn from crashes and incidents involving an ADS and can quickly respond and develop mitigation measures.

A. General guidance on ISMR implementation

129. In-Service Monitoring and Reporting (ISMR) addresses the monitoring and reporting of the in-service ADS safety performance by the manufacturer. The Monitoring refers to the overall data collection and analysis conducted by the manufactures with aim at extracting safety related information from data. The Reporting applies to occurrences which endanger

or which, if not corrected, would endanger a vehicle, its occupants or any other person, and in more terms the reporting of all occurrences relevant to the safety performance of the ADS. Annex IV provides a list of examples of these occurrences. It is expected that the ISMR will be complemented by safety investigations of (at least) critical occurrences conducted by an independent body.

130. ISMR enables the identification of unreasonable risks related to the use of an ADS on public roads and the evaluation of its safety performance during real-world operation.

131. ISMR requires ADS manufacturers to collect and analyse the safety-relevant information related to their in-service ADS' operation and report data on safety related concerns, occurrences and performance metrics to the relevant authority.

132. The ADS' safety performance remains the responsibility of the manufacturer throughout its lifetime.

133. ISMR is a mechanism to provide safety authorities with information about a manufacturer's ADS that complements information that may be gathered from other sources

1. Objectives

134. The aim of ISMR is to contribute to the improvement of road safety by ensuring that relevant information on safety is collected, processed and disseminated.

135. The ISMR aims to fulfil three main objectives:

(a) Identify safety risks related to ADS performance that need to be addressed, including instances of non-compliance with ADS safety requirements (objective 1);

(b) Support the development of the Scenario Catalogue through capturing information when the ADS does not perform safely in unanticipated situations (objective 2);

(c) Share information and recommendations to promote continuous improvement of ADS safety performance (objective 3).

136. Once there are enough ADS vehicles in-service that have encountered a sufficient range of traffic and environmental conditions then their safety needs to be evaluated. It is therefore essential that a feedback loop, facilitated by ISMR, is in place. This will provide data to assess and review the ADS manufacturer's safety case and to validate the information that was used to enable market introduction. The operational experience feedback from ISMR will allow ex-post evaluation of the regulatory requirements and validation methods, providing an indication of any issues and consequently the need for any modification.

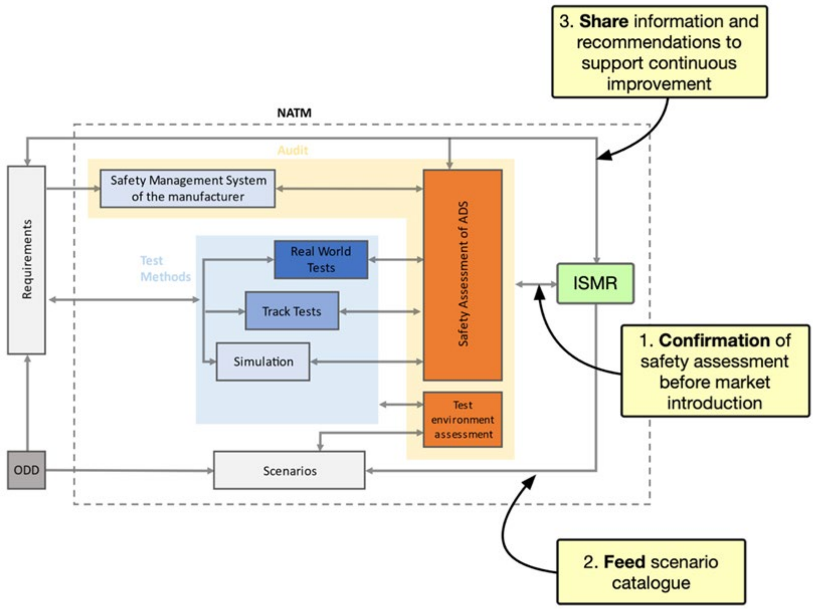
137. For example, utilising the information on ADS performance under real-world conditions could help to enhance or modify track tests. Furthermore, ISMR concerning user-interaction metrics could provide information useful for improving an ADS' HMI, its usability, and driver education.

138. Unanticipated situations, risks and hazards might be identified during real-world ADS operation, and this information could be used to develop new scenarios for the common scenario catalogue.

139. In the early phase of market introduction of ADS vehicles, it is essential that the whole community learns from safety-critical situations involving an ADS. It is important therefore that there is a mechanism that allows information from the ISMR and recommendations from its analysis to be shared with the ADS community. This will allow others to react and should lead to developments that reduce or prevent that situation from occurring in another ADS.

140. Collection, processing and dissemination of information related to ADS safety performance from the ISMR will also help to evaluate the impact of ADS on the safety of the road network.

Figure 2
ISMR integration within the multi-pillar framework



2. In Service Monitoring

141. The manufacturer and (where applicable) the fleet operator should set up a monitoring program aimed at collecting and analysing vehicle data, and data from other sources. It should provide evidence of the in-service safety performance of the ADS and confirmatory evidence of the audit results of the Safety Management System requirements established by the Audit Pillar. (Note: The in-service monitoring is intended to be applicable to all individual ADS types, not to a subset selected by the manufacturer or where applicable, by the fleet operator.)

142. The monitoring program should include a data recovery strategy, data retention strategy, data access, security and protection policy.

143. The data recovery strategy should ensure a representative collection of data to monitor the ADS in service performance.

144. The retention strategy should ensure that the dataset is retained until the corrective action and review processes are complete. In addition, the strategy should ensure the retention of the data for longer-term trend analysis (i.e. subset of the collected data).

145. The data access, security and protection policies should ensure that information access is allowed only to authorised persons and contains safeguards to ensure the security and protection of the data.

146. The data monitoring program should allow the manufacture and (where applicable) the fleet operator to:

(a) identify areas of operational risk and quantify current safety margins (e.g. in service safety performance monitoring),

(b) identify when the ADS prevents incidents/accidents (e.g. MRM, EM),

(c) identify and quantify operational risks by collecting data to characterize and analyse occurrences,

(d) use metrics and thresholds to assess safety risks and discover trends that suggest the emergence of unacceptable risks if that trend continues,

(e) put in place procedures for remedial action when an unacceptable risk is discovered or predicted by trends,

(f) confirm the in-service safety level and effectiveness of any remedial action.

147. The data monitoring program should ensure that the data analysis is performed with sufficient frequency so that remedial action can be taken promptly and in line with reporting requirements.

148. The analysis techniques should comprise the following:

(a) Routine measurements: a selection of parameters should be collected to characterise each trip and to allow a comparative analysis. These measurements should aim at identifying and monitoring emerging trends and tendencies before the trigger levels associated with exceedances are reached. (e.g. vehicle performance monitoring)

(b) Exceedance detection: a set of core ”value” should be selected to cover the main areas of interest for the ADS operation with aim at searching for deviations from vehicle performance and limits. Typically, the main areas of interest are derived from the assessment of the most significant risks before the market introduction. However, they should be continuously reviewed to reflect the current operations. (e.g., speed limits exceedance, near misses, harsh braking, etc.)

(c) Occurrence analysis: recorded data should be able to characterize and investigate all the occurrences listed in the annex IV.

(d) Statistics: Data Series should be collected to support the analysis process with additional information. These data should provide information to generate rate and trends. (e.g. driven km, operating hours).

149. The data monitoring programme should identify KPIs to assure that the monitoring is performing at an optimal level, and address any issues affecting the effectiveness of the monitoring program (e.g., data corruption or loss, or result in delayed or degraded event detection). Examples of KPIs for monitoring are trip collection rate, i.e. time between actual safety occurrence and detection of the occurrence (Date of detection of the occurrence by the In-service Monitoring – Date of the actual occurrence of the event).

(a) Vehicle data collection

150. There is regulatory work to introduce Event Data Recorder (EDR) and Data Storage System for Automated Driving (DSSAD) requirements. Until those requirements have been defined this section is only suggesting the data elements that should be collected and uploaded by the manufacturer from ADS vehicles for aggregation and processing to allow reporting of the metrics defined in the Reporting section.

(b) Other manufacturer-accessible sources of data indicative of ADS performance

151. Manufacturers may be expected to collect data relevant to typical operations such as dealer reports, customer reports, etc.

3. In Service Reporting

152. The main purpose of occurrence reporting is to identify possible improvement for the ADS safety performance, and not to attribute blame or liability.

(a) Recommended reporting by the manufacturer

153. The manufacturer should report, as required by the Authority, on both critical and non-critical occurrences, as defined in the Glossary. It is expected that two types of reports on the in-service safety performance will be produced. These are short-term and periodic.

154. Short term reporting of occurrences and safety concerns is required for matters of such safety importance that they may require the manufacturer to take remedial action, including:

(a) indications of failure to meet safety requirements

(b) critical occurrence where the ADS was involved known to the ADS manufacturer or OEM

(c) other safety-relevant performance issues

155. Short term reporting is due within one month of the manufacturer's knowledge of the matter. Short term reporting is needed to provide awareness of situations in which the ADS may be or is posing an unreasonable risk to safety in-service. Occurrences relevant to this short-term reporting are listed in Annex IV.

156. At National level, there may be further requirements for immediate reporting/notification to the authority in the event the ADS manufacturer becomes aware of a failure /defect which poses an immediate risk to public safety.

157. The manufacturer should also undertake periodic reporting of performance metrics and occurrences to the safety authority.

158. The periodic report should be delivered regularly at least every year, and should provide evidence of the in-service ADS safety performance. In particular, it should demonstrate that:

(a) no inconsistencies have been detected compared to the ADS safety performance declared prior to market introduction;

(b) the ADS respects the performance requirements set by FRAV and as evaluated in the test methods developed by VMAD;

(c) any newly discovered significant ADS safety performance issues that pose an unreasonable risk to safety have been adequately addressed and how this was achieved.

159. Annex IV provides a list of critical and non-critical occurrences aligned with FRAV's high level requirements. This represents the generic areas of interest that VMAD intends to define in greater detail. VMAD will consider both the usefulness of each suggested reporting element to the safety authorities, their capacity to review the volume of data reported, and the feasibility of storing, collecting and reporting the various elements.

160. The short term and periodic reports should be made available, as required by the Authority, in two parts:

(a) A report, that contains a summary and the information relevant to the requirements for reporting,

(b) The data underpinning the report, exchanged with the authority by means of an agreed data exchange file.

161. The authority should be informed about and agree the steps undertaken in processing the data for the report.

162. Where feasible, a consistent approach to the reporting should be developed by contracting parties, and their relevant domestic authorities.

163. The authority, where necessary, may verify the information provided and, if needed, may make recommendations to the enforcement authority and/or to the ADS manufacturer to remedy any detected conditions constituting an unreasonable risk to safety.

164. If a serious safety risk is identified, the safety authority may recommend temporary safety measures, including immediately restricting or suspending the relevant operations, and require actions to restore an acceptable level of safety.

(b) Reporting from other sources

165. The effectiveness of the ISMR pillar is determined by the availability of data on ADS safety performance. Limiting the reporting to manufacturers would also restrict the type of occurrences that may be identified by ISMR, and consequently the level of safety improvement achievable through operational experience feedback will be limited.

166. It is recommended that CPs consider extending the operational reporting mechanism to other sources (e.g. drivers, operators, users, managers, road traffic authorities …), following best practices already adopted in other transport sectors.

(c) Voluntary Reporting

167. Safety Authorities may put in place a system of voluntary reporting to collect and analyse information on observed ADS behaviours which are not required to be reported under the system of occurrences reporting set in the present Guidelines, but which are perceived by the reporter as an actual or potential hazard.

4. Collection and storage of information

168. It is recommended that a mandatory reporting system is established at national level by means of a common national database and at international level by means of a Common Central Repository.

169. Data quality and consistency should be ensured both at national and international level by establishing checking processes.

(a) National level

170. To implement the ISMR framework, Contracting Parties are recommended to designate one or more competent authorities to put in place a mechanism to collect, evaluate, process and store occurrences reported in accordance with ISMR principles.

171. The safety authority/ies at national level should be responsible for collecting and assessing the data and for deriving and sharing safety recommendations. It (They) should manage the safety-related information stored in the national database and share that information with other competent authorities. The safety authority is also in charge of issuing an annual report summarizing the level of ADS safety and providing an overall safety assessment and action plan. The annual report should be submitted to WP29.

172. Short term and periodic reports should be stored within the common national database. Safety recommendations should also be stored in the common national database and made accessible to the relevant stakeholders.

173. Safety authorities should transfer safety recommendations and annual reports to the Common Central Repository.

(b) International level

174. WP29 provides a suitable international context for exchanges between Contracting Parties and for defining the guiding principles on the ISMR framework implementation.

175. It is recommended that WP.29 establishes a proper management system of the Common Central Repository. It should cover accessibility and dissemination of information, data protection where needed, data evaluation and annual reporting. The technical protocols for transferring all safety recommendations to the Common Central Repository should also be established.

176. Clear guidance on the standardized approach to ISMR, including the harmonisation of the data entry process, should be organized by WP.29 at international level by providing guidelines, workshops and appropriate training.

5. Occurrences Investigations

177. It is recommended that Contracting Parties designate one competent body responsible for conducting the investigations of accidents, incidents and any other relevant event in their countries according to its investigation mandate. The body may be an existing transportation safety investigative agency responsible for investigating transportation accidents.

178. It is desirable for this body to be independent in its organisation, legal structure and decision-making from any interested party, including other entitled regulatory body, other national bodies in charge of investigating liability aspects of crashes or in charge of the collection and storage of information reported by manufacturers.

179. In case of accidents/incidents an investigation report should be produced. It should be produced and made available in the shortest possible time after the date of the occurrence to all parties involved. It should where appropriate, contain safety recommendations.

180. A periodic report should be produced and shared regularly at least every year, or more frequently if relevant. It should provide information about the investigations carried out in the preceding year and the safety recommendations that were issued.

6. Exchange of Information

181. It is recommended that WP29 promotes and facilitates a broader exchange of information and the dissemination of safety recommendations among the Contracting Parties with the aim of improving safety.

182. Safety Authorities should participate regularly in the exchange and analysis of information contained in the Common Central Repository.

183. It is recommended that Safety Authorities participate in an exchange of information by making all relevant safety-related information available to the other competent authorities.

184. The exchange of relevant information among involved Contracting Parties / Authorities should be required in case of accidents/incidents investigations.

185. The dissemination of information should be limited to what is strictly required for the purpose of its users, in order to ensure appropriate confidentiality of that information.

7. Protection of information

186. Given the sensitive nature of safety-related information, the protection of its source and the confidence and trust of the reporters should be guaranteed. To protect the sensitivity of the information, it is recommended that it is only used for safety related activities and not for any other purpose.

187. Security measures need to be in place to protect the confidentiality of information that is shared. For example, the security measures and protocols should ensure that no personal details are ever recorded in the databases either at national or international level and that relevant protections for trade secrets and confidential business information be observed.

188. Without prejudice to the applicable national law, it is recommended that Safety Authorities refrain from instituting proceedings in respect of unpremeditated or inadvertent infringements of the law that come to their attention only because they have been reported under the ISMR occurrence-reporting scheme, except in cases of gross negligence.

189. In accordance with the procedures defined in their national laws and practices, Safety Authorities should ensure that employees who report incidents of which they may have knowledge are not subjected to any prejudice by their employer.

XI. NATM Pillars/Element Interaction

190. The goal of the NATM guidelines document is to assess the safety of an ADS in a manner that is as repeatable, objective and evidence based as possible, whilst remaining technology neutral and flexible enough to foster ongoing innovation in the automotive industry.

191. The overall purpose of the NATM is to assess, based on the safety requirements, whether the ADS is able to cope with occurrences that may be encountered in the real world. In particular, by looking at scenarios linked to road users' behaviour/environmental conditions in Traffic scenarios as well as scenarios linked to driver behaviour (e.g. HMI) and ADS failures.

192. As previously noted, the multi-pillar approach recognizes that the safety of an ADS cannot be reliably assessed/validated using only one of the pillars. Each of the aforementioned testing methodologies possesses its own strengths and limitations, such as differing levels of environmental control, environmental fidelity, and scalability, which should be considered accordingly.

193. It is important to note that a single assessment or test method may not be enough to assess whether the ADS is able to cope with all occurrences that may be encountered in the real world.

194. For instance, while real-world testing provides a high degree of environmental fidelity, a scenario-based testing methodology using only real-world testing could be costly, time-consuming, difficult to replicate, and pose safety risks. Consequently, track testing may be more appropriate methods to run higher risk scenarios without exposing other road users to potential harm. Further, test scenarios can also be more easily replicated in a closed track environment compared to the real-world. That said, test track scenarios can be potentially difficult to develop and implement, especially if there are numerous or complex scenario*s*, involving a variety of scenario elements.

195. Consideration should be given to the fact that simulation/virtual testing, by contrast, can be more scalable, cost-effective, safe, and efficient compared to track or real-world testing, allowing a test administrator to safely and easily create a wide range of scenarios, including complex scenarios, where a diverse range of elements are examined. However, simulations may be of a lower fidelity than the other methodologies. Simulation software may also vary in quality and tests could be difficult to replicate across different simulation platforms.

196. In-service monitoring and reporting should be used to confirm the pre-deployment safety assessment and fill the gaps between safety validation through virtual/physical testing and real-life conditions. Evaluation of in-service performance should also be used to update the scenario database with new scenarios derived from the increased deployment of ADS. Finally, the feedback from operational experience can support ex-post evaluation of regulatory requirements.

197. In addition to the respective strengths and weakness of each test pillar, the nature of the safety requirements being assessed will also inform what pillars are used:

(a) For instance: the most appropriate method to assess an ADS's overall system safety prior to market introduction may be the audit pillar, using a systematic approach to perform a risk analysis. The audit could include information such as safety by design confirmed validation outputs as well as analysis of data collected in the field by the manufacturer.

(b) Virtual testing may be more suitable when there is a need to vary test parameters and a large number of tests need to be carried out to support efficient scenario coverage (e.g., for path planning and control, or assessing perception quality with pre-recorded sensor data).

(c) Track tests may be best suited for when the performance of an ADS can be assessed in a discrete number of physical tests, and the assessment would benefit from higher levels of fidelity (e.g., for HMI or fall back, critical traffic situations).

(d) Real-world testing may be more suitable where the scenario may not be precisely represented virtually or on a test track (e.g., interactions with other road-users and perception quality may be assessed through real world evaluation).

(e) In-service monitoring and reporting of field data represent the best way to confirm the safety performance of an ADS in the field after market introduction over a wide variety of real driving traffic and environmental conditions.

198. Given these considerations, it should be noted that the sequence and composition of test pillars used to assess each safety requirement may vary. While some testing might follow a logical sequence from simulation to track and then to real world testing, there may be deviations depending on the specific safety requirement being tested.

199. It is therefore necessary for the NATM pillars to be used together to produce an efficient, comprehensive, and cohesive process, considering their strengths and limitations. The methods should complement one another, avoiding excessive overlaps or redundancy to ensure an efficient and effective validation strategy.

200. As previously noted, the NATM pillars not only include the three aforementioned test methods but also an aggregated analysis (e.g., an audit/assessment /in service monitoring/reporting pillar). Whereas the test methods will assess the safety of the ADS, the audit/assessment pillar will serve to assess the safety of the ADS as well as the robustness of organizational processes/strategies. Elements of the audit are:

(a) Assessment of the robustness of safety management system.

(b) Assessment of the (identified) hazards and risks for the system.

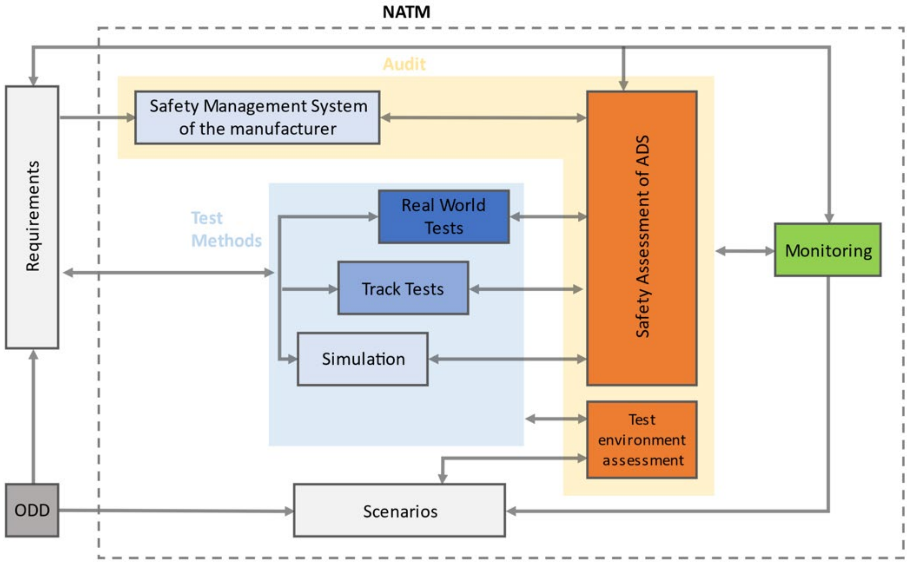
(c) Assessment of the Verification strategy (e.g. verification plan and matrix) that describe the validation strategy and the integrated use of the pillars to achieve the adequate coverage.

(d) Assessment of the level of compliance with requirements achieved through an integrated use of all pillars, including consistency between the outcomes of one pillar as input for another pillar (forward and backward) and adequate use of scenarios. This level of compliance concerns both new vehicles as vehicles in use.

(e) The audit/assessment phase also incorporate results from the Simulation, Track test and Real-World tests carried out by the manufacturer.

201. Figure 3 provides a diagram that outlines how the pillars, scenarios, and safety requirements (developed by FRAV) will interact. Further examination of each of these elements follows in the subsequent sections of this document.

Figure 3
Relationship between VMAD Pillars, Scenarios and FRAV Safety Requirements



XII. VMAD NATM- FRAV Integration

202. This document contains the description of a generic validation method. Likewise, FRAV (Functional Requirements for Automated Vehicles) is developing generic requirements for the product to be validated. There is a clear relation between these two developments: functional requirements may affect the detailed validation requirement and vice versa. Furthermore, validation requirement may result in input for functional requirements.

203. So far, FRAV has delivered a list of high level safety requirements. In detailing the functional requirements, the possible impact for validation methods will have to be checked. This process is managed by including representatives of both informal working groups in each other's meetings.

204. As the safety requirements and technical aspects of each of the pillars are further developed, each of these sections will be updated to include additional detail. To provide further context, this section will also include examples of how the NATM pillars can be applied to certain functional capabilities of an ADS (e.g., highway driving, which is described further in Annex II.) based on the established safety requirements. FRAV and VMAD will continue to engage to develop and update functional requirements and the technical aspects

of each pillar as necessary. This is key to ensuring safety guidance is updated as ADS technologies evolve.

Annex I

Glossary of Terms and Definitions

"Abstraction" is the process of selecting the essential aspects of a source system or referent system to be represented in a model or simulation, while ignoring those aspects not relevant. Any modelling abstraction carries with it the assumption that it should not significantly affect the intended uses of the simulation tool.

[*"Abstract Scenario"*: A formalized, machine readable, declarative description of the scenario, derived from functional scenario. The declarative specification on the abstract level enables highlighting of the relevant aspects of the scenario while focusing on efficient description of relations (Cause-effect).]

"Automated Driving System (ADS)" means the vehicle hardware and software that are collectively capable of performing the entire Dynamic Driving Task (DDT) on a sustained basis.

"ADS feature" means an application of an ADS designed specifically for use within an Operation Design Domain (ODD).

"ADS function" means an application of ADS hardware and software designed to perform a specific portion of the DDT.

"Closed Loop Testing" means a virtual environment that does take the actions of the element-in-the loop into account. Simulated objects respond to the actions of the system (e.g. system interacting with a traffic model).

"Concrete Scenarios": [A scenario depicted with explicit parameters values, describing physical attributes.] Concrete scenarios are established by selecting specific values for each element. This step ensures that a specific test scenario is reproducible. In addition, for each logical scenario with continuous ranges, any number of concrete scenarios can be developed, helping to ensure a vehicle is exposed to a wide variety of situations.

"Complex Scenarios" means a traffic scenario containing one or more situations that involve a large number of other road users, unlikely road infrastructure, or abnormal geographic/environmental conditions.

"Critical Scenarios" means a traffic scenario containing a situation in which the ADS needs to perform an emergency maneuver in order to avoid/mitigate a potential collision, or react to a system failure.

"Deterministic" is a term describing a system whose time evolution can be predicted exactly and a given set of input stimuli will always produce the same output.

"Driver-In-the-Loop" *(DIL*) is typically conducted in a driving simulator used for testing the human–automation interaction design. DIL has components for the driver to operate and communicate with the virtual environment.

"Dynamic driving task (DDT)" means all of the real-time operational and tactical ADS functions required to operate the ADS-equipped vehicle in on-road traffic.

- The DDT excludes strategic functions such as trip scheduling and selection of destinations and waypoints.
- The DDT functions can be logically grouped under three main categories:

(a) Sensing and Perception, including;

- Monitoring the driving environment via object and event detection, recognition, and classification,

- Perceiving other vehicles and road users, the roadway and its fixtures, objects in the vehicle's driving environment and relevant environmental conditions,
- Sensing the ODD boundaries, if any, of the ADS feature,
- Positional awareness

(b) Planning and Decision, including;

- Prediction of actions of other road users.
- Response preparation.
- Maneuver planning

(c) Vehicle Control, including;

- Object and event response execution.
- Lateral vehicle motion control.
- Longitudinal vehicle motion control.
- Enhancing conspicuity via lighting and signaling

"Edge Case" is a rare situation that may require specific design attention for it to be dealt with by the ADS in a reasonable and safe way if warranted by the possible severity and likely frequency within the ODD of the ADS. The quantification of "rare" is relative, and generally refers to situations or conditions that will occur often enough in a full-scale deployed fleet to be a problem but may have not been captured in the design process. Edge cases can be individual unexpected events, such as the appearance of a unique road sign or an unexpected animal type on a highway

"Functional Scenario": [A scenario described in natural language on a conceptional level, in general without specific physical values. These are scenarios / Scenarios] with the highest level of abstraction, outlining the core concept of the scenario, such as a basic description of the ego vehicle's actions; the interactions of the ego vehicle with other road users and objects; and other elements that compose the scenario (e.g. environmental conditions etc.). This approach uses accessible language to describe the situation and its corresponding elements. For the scenario catalogue, such an accessible (i.e., natural and non-technical) language needs to be standardised to ensure common understanding between different ADS stakeholders about the scenarios.

"*Hardware-In-the-Loop*" (HIL) involves the final hardware of a specific vehicle sub-system running the final software with input and output connected to a simulation environment to perform virtual testing. HIL testing provides a way of replicating sensors, actuators and mechanical components in a way that connects all the I/O of the Electronic Control Units (ECU) being tested, long before the final system is integrated.

"Logical Scenario": [A scenario described with the inclusion of parameters, where the values of some of the parameters are defined as ranges.] Building off the elements identified within the functional scenario, developers generate a logical scenario by selecting value ranges or probability distributions for each element within a scenario (e.g., the possible width of a lane in meters).

"Model" is a description or representation of a system, entity, phenomenon, or process.

"Model calibration" is the process of adjusting numerical or modelling parameters in the model to improve agreement with a referent.

"Model-In-the-Loop" (MIL) is an approach which allows quick algorithmic development without involving dedicated hardware. Usually, this level of development involves high-level abstraction software frameworks running on general-purpose computing systems.

"Model Parameter" are numerical values used to support characterizing a system functionality. A model parameter has a value that cannot be observed directly in the real

world but that must be inferred from data collected in the real world (in the model calibration phase).

"Nominal Scenarios" means a traffic scenario containing situations that reflect regular and non-critical driving manoeuvres.

"*Occurrence*" refers to any safety-related event involving a vehicle equipped with an ADS. For reporting, two different categories of occurrences are defined.

"*Non-critical Occurrence*" means an operational interruption, defect, fault or other circumstance that has or may have influenced ADS safety but has not resulted in an accident or serious incident. This category includes for example minor incidents, safety degradation not preventing normal operation, emergency/complex manoeuvres to prevent a collision, and more generally all occurrences relevant to the safety performance of the in-service ADS (like transfer of control, interaction with remote operator, etc.).

"Critical Occurrence" means an occurrence in which the ADS is engaged at the time of the event and:

(a) at least one person suffers an injury that requires medical attention as a result of being in the vehicle or being involved in the event;

(b) the ADS vehicle, other vehicles or stationary objects sustain physical damage that exceeds a certain threshold;

(c) any vehicle involved in the event experiences an airbag deployment

"*Operational Design Domain (ODD)*" means the operating conditions under which an ADS feature is specifically designed to function.

"*ODD exit*" means:

(a) the presence of one or more ODD conditions outside the limits defined for use of the ADS feature, and/or

(b) the absence of one or more conditions required to fulfil the ODD conditions of the ADS feature.

"Open Loop Testing" means a virtual environment that does not take the actions of the element-in-the loop into account (e.g. system interacting with a recorded traffic situation).

"Probabilistic" is a term pertaining to non-deterministic events, the outcomes of which are described by a measure of likelihood.

"Proving Ground or test-track" is a physical testing facility closed to the traffic where the performance of an ADS can be investigated on the real vehicle. Traffic agents can be introduced via sensor stimulation or via dummy devices positioned on the track.

"Sensor Stimulation" is a technique whereby artificially generated signals are provided to the element under testing in order to trigger it to produce the result required for verification of the real world, training, maintenance, or for research and development.

"Simulation" is the imitation of the operation of a real world process or system over time.

"Simulation toolchain" is a combination of simulation tools that are used to support the validation of an ADS.

"Software-In-the-Loop" (SIL) is where the implementation of the developed model will be evaluated on general-purpose computing systems. This step can use a complete software implementation very close to the final one. SIL testing is used to describe a test methodology, where executable code such as algorithms (or even an entire controller strategy), is tested within a modelling environment that can help prove or test the software.

*"Stochasti*c" means a process involving or containing a random variable or variables. Pertaining to chance or probability.

"Test case specification" are the detailed specifications of what must be done by the tester to prepare for the test.

"Test methods" is a structured approach to consistently derive knowledge about the ADS by means for executing tests, e.g. virtual testing in simulated environments, physical, structured testing in controlled test facility environments, and real world on-road conditions.

"Traffic scenario" (or scenario for short) is a sequence or combination of situations used to assess the safety requirements for an ADS. Scenarios include a driving maneuver or sequence of driving maneuvers. Scenarios can also involve a wide range of elements, such as some or all portions of the DDT; different roadway layouts; different types of road users and objects exhibiting static or diverse dynamic behaviours; and, diverse environmental conditions (among many other factors).

"*Transfer of Control (TOC)"* means a transfer of dynamic control of the vehicle from the ADS to the ADS vehicle user.

"*TOC request*" means a warning issued by the ADS to the fallback user that the latter is needed to engage in dynamic control of the vehicle.

"*TOC response*" means the fallback user engagement in the dynamic control of the vehicle pursuant to a TOC request.

"Validation of the simulation model" is the process of determining the degree to which a simulation model is an accurate representation of the real world from the perspective of the intended uses of the tool.

"Vehicle -In-the-Loop" (VIL) is a fusion environment of a real testing vehicle in the real-world and a virtual environment. It can reflect vehicle dynamics at the same level as the real-world and it can be operated on a vehicle test bed or on a test track.

"Verification of the simulation model" is the process of determining the extent to which a simulation model or a virtual testing tool is compliant with its requirements and specifications as detailed in its conceptual models, mathematical models, or other constructs.

"Virtual testing" is the process of testing a system using one or more simulation models.

Annex II

Functional Scenarios for divided highway application

Contents

I. Introduction

This text is a synthesis of various recent elaborations of Traffic scenarios, with the designated purpose to create a functional scenario list for ADS in motorway use-case. It is envisaged that some logical scenarios and/or some possible ways of their description, as agreed in the continuous discussion, will also be included in this text. ODD range: highways with up to 130 km/h and lane changes allowed.

II. Inputs to this proposal

(a) Present UN ALKS regulation (R157)

(b) The Netherlands (TNO) Scenario Categories V1.7

(c) SAFE (Fortellix) scenario library **Error! Reference source not found.**

(d) Japan Crash scenarios

(e) China functional scenario proposal (CATARC)

(f) JRC own elaborations

(g) Germany (IGLAD) catalogue of conflict types

Inputs provided by JP, NL, SAFE, CN were submitted for consideration and discussion during the VMAD SG1 meeting held on 10 December, proposal from DE submitted on 16 December 2020.

III. Building blocks of functional scenarios

As described previously in the Scenario Catalogue section, functional scenarios can cover several aspects (e.g. road geometry at different abstraction levels, ego-vehicle behaviour, moving/stable objects).

Additional aspects that are not covered by functional scenarios (e.g. speeds, accelerations, positions, environmental conditions, failures, miscommunications, road geometries at more detailed levels) should be covered by logical scenario.

Since classification of aspects to functional and logical scenarios (i.e. “which aspects should be considered in functional scenarios” and “which aspects should be considered in logical scenarios”) has not yet been discussed and agreed, the classification in this document is initial version and will be updated through discussion.

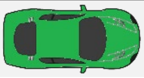
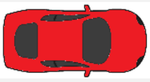
IV. Coverage

Collisions always occur with other vehicles/objects (assuming that they can operate properly when there are no other vehicles/objects). The 24 functional scenarios in the figure described in section “2. Interaction with other vehicles” under Nominal Driving can cover all interactions between other vehicles/objects and ego vehicle. These scenarios can cover collision with other vehicles/objects appropriately.

As described above., factors not covered in the proposed functional scenarios (e.g. initial speed of ego vehicle, size, initial position, initial speed, acceleration of other vehicles/objects), perception factor (e.g. weather, brightness, blind spot, false positive factor, blinkers of other vehicles) and vehicle stability factors (e.g. curve, slope, road surface μ, wind, etc.) can be described with parameters in logical scenarios.

As previously mentioned, it is anticipated that future iterations of Annex II will also incorporate scenarios with lower levels of abstraction (e.g. logical scenarios and potential approaches for describing them). Functional scenarios should be added when agreement is reached between SG1 and VMAD-IWG.

V. Symbols used in this document

ICON	DESCRIPTION
	Ego vehicle
	Lead vehicle
	Other vehicles part of the scenario
	Impassable object on intended path
	Passable object on intended path

VI. A list of possible scenarios for L3 Highway Chauffeur ADS

Input matrix from VMAD-SG1 participants:

Scenario family		*Sub-scenario*	*Japan crash scenarios*	*The Netherlands (TNO)*	*SAFE scenario library*	*China functional scenarios*	*Conflict Type*
1. Nominal driving	1-1. Perform lane keeping	a. Driving straight		X	X	X	X
		b. Manoeuvring a bend		X	X	X	X
2. Interaction with other vehicles/ objects	2-1. Perform lane change	a. Ego vehicle performing lane change with vehicle behind	X	X			X
		b. Merging at highway entry	X		X	X	X
		c. Merging at lane end	X		X		X
		d. Merging into an occupied lane	X	X			X
	2-2.Critical (Emergency) braking scenarios during lane keeping	e. Impassable object on intended path	X	X	X		X
		f. Passable object on intended path	X	X		X	X
		g. Lead vehicle braking	X	X	X	X	X
		h. Approaching slower/stopped LV	X	X	X	X	X
		i. Cut-in in front of the ego vehicle	X	X	X	X	X
		j. Cut-out in front of the ego vehicle	X	X	X	X	X

Scenario family		*Sub-scenario*	*Japan crash scenarios*	*The Netherlands (TNO)*	*SAFE scenario library*	*China functional scenarios*	*Conflict Type*
		k. Detect and respond to swerving vehicles	X	X	X		X
3. Detect and response to traffic rules and road furniture		a. Speed limit sign			X	X	
		b. Signal lights				X	X
		c. Drive through tunnel				X	
		d. Toll				X	
		e. Conventional obstacles				X	X
4.Country specific road geometry		a. Intercepter			X		
5. Unusual situation		a. Wrong way driver (oncoming)		X			X

Notes to the inputs from VMAD SG1 members:

- China (CATARC): This is a list cut from a general catalogue describing different ODDs, like “General road”, “City expressway” or “The highway” and their test items, like “speed limit sign”, “lane line”, “toll station”, etc. The functional scenarios proposed below in this document are much more generic than the ones proposed by China, so they form a subset of this list. For example, China proposal: “toll station” on the road or “conventional obstacles” can be in line with “impassable object on intended path” from this scenario list.
- The Netherlands (TNO): a very thorough scenario catalogue containing much more scenarios than needed for the highway use case. Terminology and descriptions worked out fully. Scenarios can be created using a combination of tags from the different layers.
- Japan: crash scenarios, scenarios only containing interaction with other vehicles. They describe different road geometries and possible other vehicle positions around ego. All other parameters considered as features (acceleration – deceleration, lane change – lane keeping, etc.).
- SAFE: a list of scenarios sometimes with very concrete examples, sometimes more generic approach. There is a different scenario for passing by slowly moving vehicles in the adjacent lane and a different one for passing by standing vehicles, but handles LV following as one scenario.
- Conflict Type: a list of “conflict types” used i.a. by accident investigators to sort scenarios, leading to accidents on road to different groups. These conflict types can be sorted into conflicts with or without influence of other road user. Uses different symbols than other documents for the description of a scenario or situation (mainly different kinds of arrows). Separates left and right hand traffic. Contains 251 scenario types, structured in seven larger types of conflicts, like: “longitudinal traffic” or “pedestrian crossing the road”.

Note: "emphasized scenario parameters" and "tested parameters" in this paragraph are some examples of parameters. Other parameters may be essential for the validation testing.

A. Nominal driving (Perform lane keeping)

1. Nominal driving (Perform lane keeping)

Note: lane keeping is addressed in current UN-Regulation for ALKS No. 157 up to 60 km/h. As a functional scenario, lane keeping can be sorted into two groups depending on road geometry. It can also be sorted into more groups depends on the lane that the vehicle is in: centre, side, middle, etc.

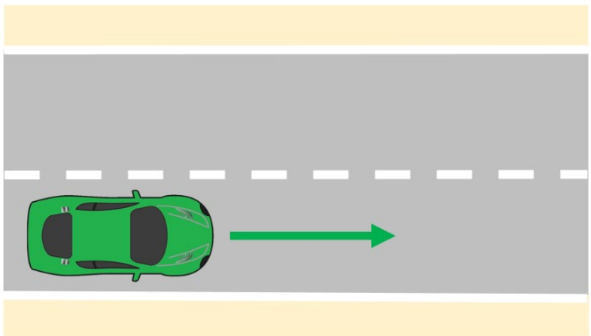
(a) Driving straight

(a) Without LV

(b) With LV

(c) With other vehicles in adjacent lanes (moving or stopped)

Figure 1
Schematic representation of driving straight



General description:

The ego vehicle is driving on a straight road. The aim of this scenario is to test the lane keeping ability of the vehicle under normal or demanding conditions and parameters [1,2,4].

Emphasized scenario parameters: ego speed demand (road rules), lane width, LV speed profile (if present), layout and speed profile of other vehicles (if present).

Tested parameters: deviation from lane centre (nominal value and distribution), deviation from desired speed, obeying to speed changes, temporal modifications, distance between ego and LV (if present), reaction to other vehicles etc.

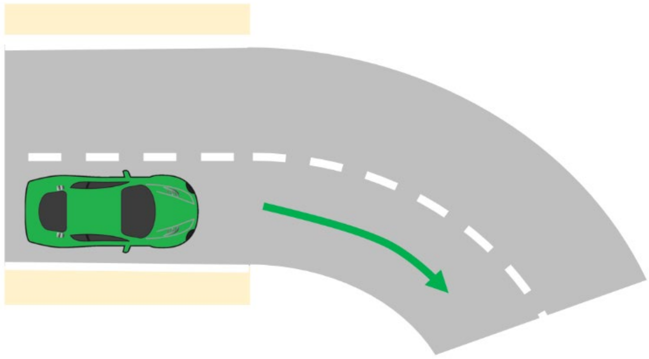
(b) Manoeuvring a bend (right curve and left curve)

(a) Without LV

(b) With LV

(c) With other vehicles in adjacent lanes (moving or stopped)

Figure 2
Schematic representation of manoeuvring a bend



General description:

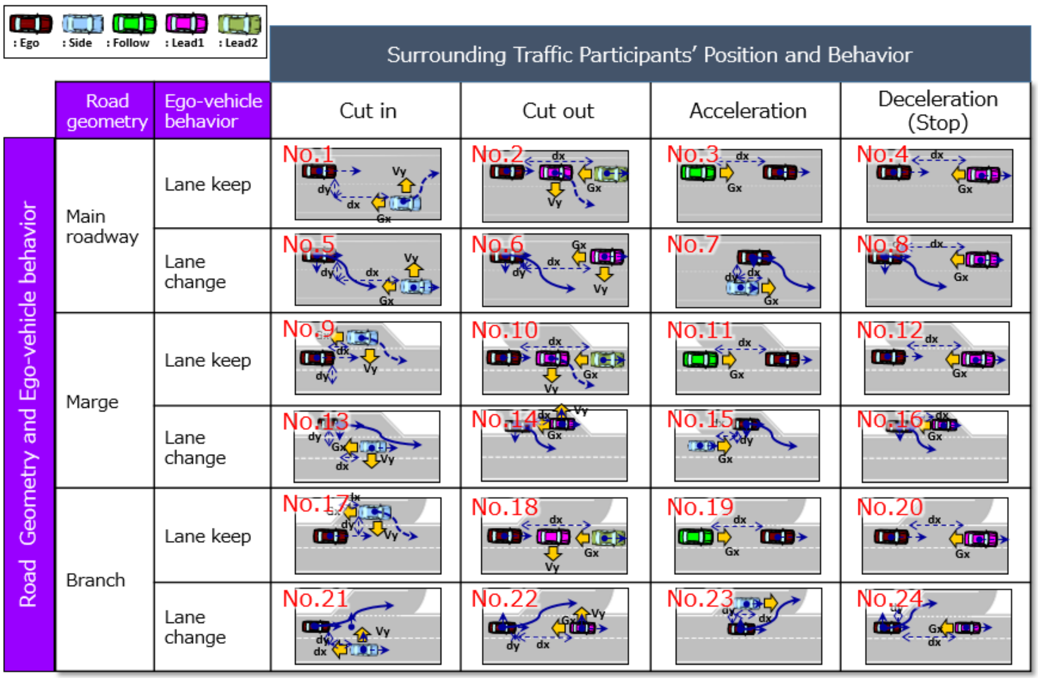
The ego vehicle is driving on a curved road. The aim of this scenario is to test if the vehicle is able to handle the road curvatures specified as part of the ODD [1,2,4].

Emphasized scenario parameters: ego speed demand (road rules), lane width, LV speed profile (if present), layout and speed profile of other vehicles (if present).

Tested parameters: deviation from lane centre (nominal value and distribution), deviation from desired speed, obeying to speed changes, temporal modifications, distance between ego and LV (if present), distance to other vehicles etc.

B. Interaction with other vehicles/objects

The 24 scenarios below can cover the interaction with other vehicles driving in the same direction on the same or adjacent lanes.



In the 12 scenarios in which the ego vehicle performs lane change, the vehicle closest to the ego vehicle may not be necessarily in the same lane or an adjacent lane to the ego vehicle. It may be 2 lanes over from the ego vehicle, and even in such cases, the vehicle has to be detected by the ego vehicle because they can interact with one another if both change lanes. To describe these cases in the 12 scenarios properly, some parameters should be included such as "number of lanes", "lane of ego vehicle" and "relative position between ego and other vehicle". The examples of "main road case" are shown below. Other cases in "merged road" and "branched road" should be considered too.

	Main road 2 lanes (1st lane / 2nd lane)			Main road 3 lanes (1st lane / 2nd lane / 3rd lane)		
	Forward	Parallel running	Rear	Forward	Parallel running	Rear
No.5 LC in the opposite direction						
No.6 LC in the same direction						
No.7 Acceleration						
No.8 Deceleration						

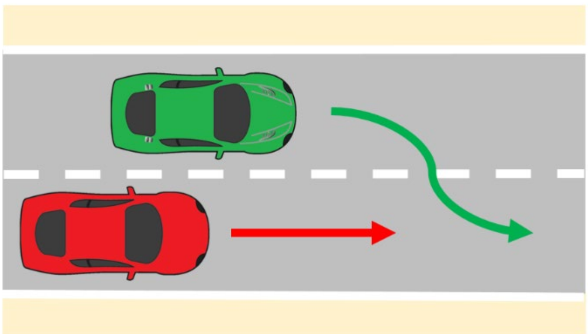
1. **Perform lane change**

Note: LC scenarios are complicated by the fact that the ADS cannot be forced to make a lane change. In addition, lane change functionality and principles shall be defined in a later stage (like technical requirements, definitions, activation criteria, indication of lane change, etc.).

Lane changes can be grouped based on the number of vehicles in the target lane. If there is enough space to execute the lane change, there is no need to cooperate with other vehicles. If the target lane is occupied by other traffic participants, than the ego vehicle has to adapt to the other participants and perform merging.

(a) **Ego vehicle performing lane change with vehicle behind**

Figure 3
Schematic representation of a lane change



General description:

In an adjacent lane, another vehicle is driving in the same direction as the ego vehicle. The intention of the ego vehicle is, to perform a lane change to the lane in which the other player is driving [1,3].

Emphasized scenario parameters: time of lane change, ego speed demand (road rules), lane width, LV speed profile (if present), layout and speed profile of other vehicles (if present).

Tested parameters: deviation from lane centres (nominal value, overshoot), time of lane change (lateral velocity of ego), distance between ego and LV (if present), distance to other vehicles, etc.

(b) **Merging at highway entry**

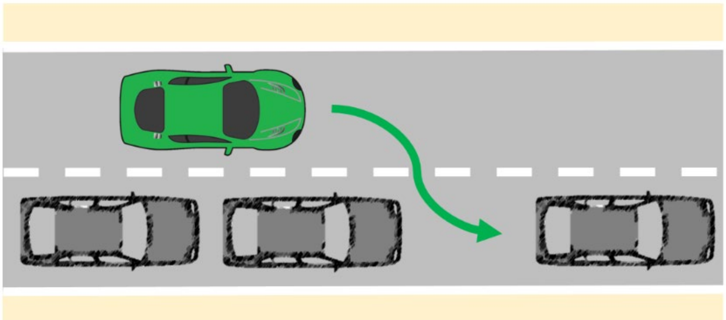
No description provided

(c) **Merging at lane end**

No description provided

(d) **Merging into an occupied lane**

Figure 4
Schematic representation of merging



General description:

Other vehicles occupy the lane adjacent to the ego lane. The ego vehicle intends to perform a lane change to the lane in which the other vehicles are driving [1-4]. According to road geometry, speed, number and layout of other vehicles, the difficulty of the scenario changes.

Emphasized scenario parameters: road layout, layout and speed profile of other vehicles (if present), ego speed (road rules), lane width etc.

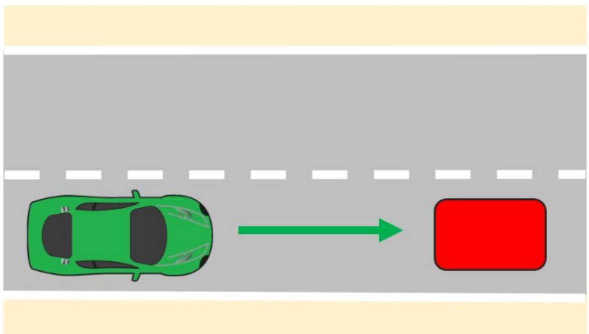
Tested parameters: distance to other vehicles, time of lane change (lateral velocity of ego).

2. **Critical (Emergency) braking scenarios during lane keeping**

Note: In this family of scenarios a couple critical functional scenarios are present. It can be noticed in the input matrix of SG1 as well, these are scenarios that nearly every participant highlighted in the input documents.

(a) **Impassable object on intended path (Including other cars and VRUs)**

Figure 5
Schematic representation of an impassable object



General description:

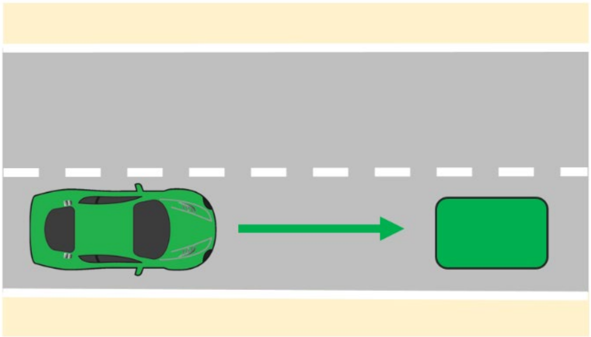
The ego vehicle is driving on a road with an impassable object in the ego lane. The objective of the ego vehicle is to continue driving straight. The ego vehicle needs to react [1,2]. Depending on the velocity of the ego vehicle, the severity of the scenario is changing.

Emphasized scenario parameters: road layout (visibility of the object on the path), layout and speed profile of other vehicles (if present), ego velocity.

Tested parameters: reaction of ego (lane change/braking), distance to object, lateral velocity of ego (if changing lane) etc.

(b) Passable object on intended path (e.g. manhole lid)

Figure 6
Schematic representation of a passable object



General description:

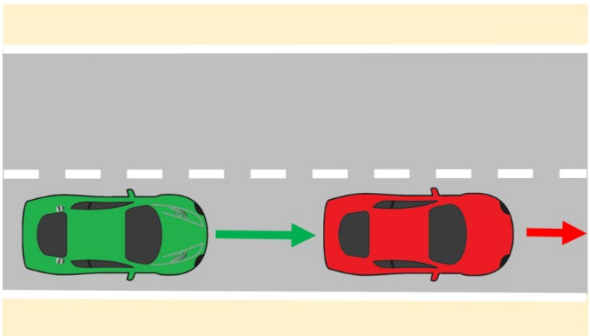
The ego vehicle is driving on a road with a passable object in the ego lane, e.g., a manhole lid or a small branch. The objective of the ego vehicle is to continue driving straight. The ego vehicle needs to react [1,4]. Depending on the velocity of the ego vehicle, the difficulty of the scenario is changing.

Emphasized scenario parameters: road layout (visibility of the object on the path), layout and speed profile of other vehicles (if present), ego velocity.

Tested parameters: reaction of ego (false positive, lane change/braking), distance to object, lateral velocity of ego (if changing lane), etc.

(c) Lead vehicle braking

Figure 7
Schematic representation of lead vehicle braking



General description:

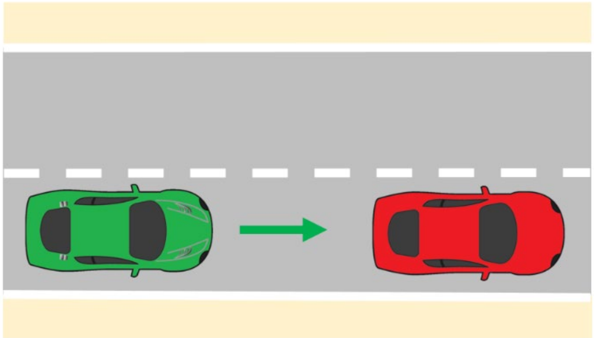
The ego vehicle is following a LV. The LV brakes, the ego vehicle has to adapt its speed in order to stay at a safe distance from the lead vehicle [1-4].

Emphasized scenario parameters: ego velocity (road rules), LV speed profile (deceleration), layout and speed profile of other vehicles (if present).

Tested parameters: distance between ego and LV, reaction to other vehicles in adjacent lanes, etc.

(d) Approaching slower/stopped LV

Figure 8
Schematic representation of approaching stopped lead vehicle



General description:

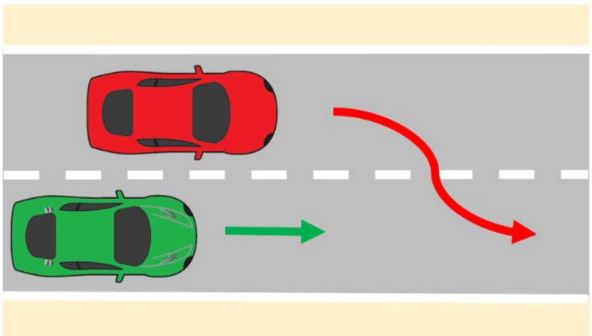
LV is driving in front of the ego vehicle at a slower speed. The ego vehicle might brake or perform a lane change to avoid a collision [1-4]. According to the speed of the LV and ego vehicle, the severity of this scenario can be assessed.

Emphasized scenario parameters: ego velocity (road rules), LV speed profile (deceleration), layout and speed profile of other vehicles (if present).

Tested parameters: distance between ego and LV, reaction to other vehicles in adjacent lanes, etc.

(e) Cut-in in front of the ego vehicle

Figure 9
Schematic representation of cut-in



General description:

Another vehicle is driving in the same direction as the ego vehicle in an adjacent lane. The other vehicle makes a lane change, such that is becomes the LV from the ego vehicle's perspective [1-4]. Depending on the distance and lateral velocity of the LV, the severity of the cut-in manoeuvre changes.

Emphasized scenario parameters: LV lateral speed, distance to LV, ego velocity, lane width, layout and speed profile of other vehicles (if present).

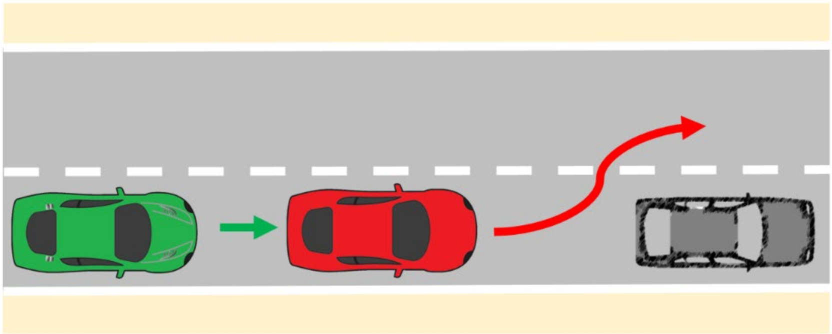
Tested parameters: distance between ego and LV, distance to other vehicles, etc.

(f) Cut-out in front of the ego vehicle

(a) Cut-out to highway exit

(b) Cut-out on highway lanes

Figure 10
Schematic representation of cut-out



General description:

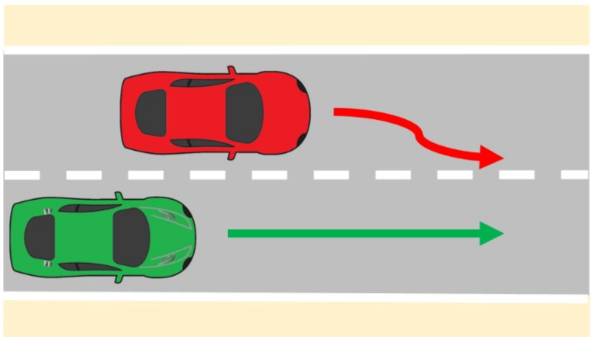
LV is driving in the same direction as the ego vehicle in front of the ego vehicle. The LV makes a lane change, such that it will no longer be the ego vehicle's LV [1-4]. In order to test the behaviour of the ego vehicle, an obstacle is present in the ego lane in front of the ego vehicle. Depending on the velocity of the ego vehicle and the lateral velocity of the LV, the difficulty of this scenario changes.

Emphasized scenario parameters: LV lateral speed, distance to LV, ego velocity, lane width, layout and speed profile of other vehicles (if present).

Tested parameters: distance between ego and obstacle, distance to other vehicles etc.

(g) Detect and respond to swerving vehicles

Figure 11
Schematic representation of a swerving vehicle



General description:

Another vehicle is driving in the same direction as the ego vehicle in an adjacent lane. The other vehicle swerves towards the ego vehicle's lane [1-3].

Emphasized scenario parameters: lateral speed of other vehicle, ego velocity, lane width, layout and speed profile of other vehicles (if present).

Tested parameters: distance between ego and swerving vehicle, distance to other vehicles, etc.

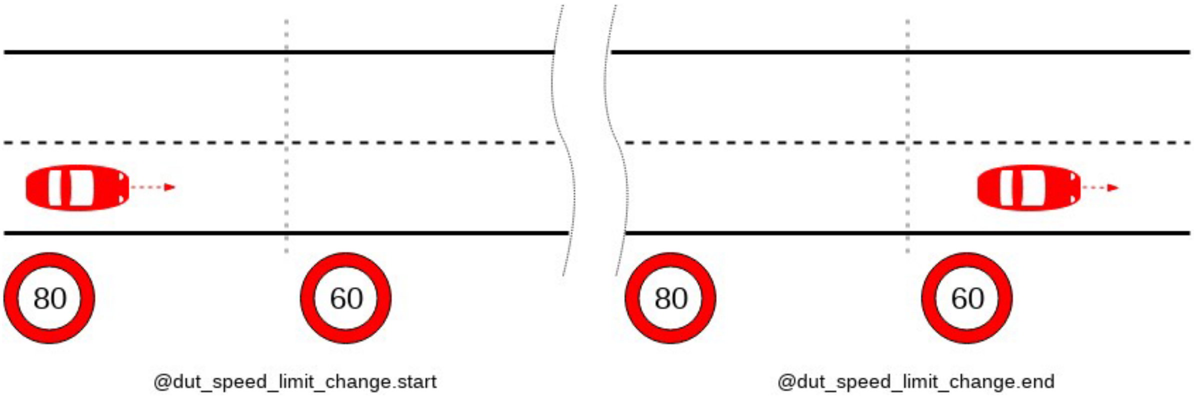
C. Detect and respond to traffic rules and road furniture

Note: These scenarios are implicitly present in nearly every document, but sometimes are treated as special road furniture. It should be considered that these scenarios can be occurred simultaneously with other scenarios. It should be also noted that traffic rules are different from different countries or regions.

(a) Speed limit sign

This scenario challenges the ego vehicle to respond appropriately to speed limit changes by decelerating when entering a lower speed zone and accelerating when entering a higher speed zone. In the example shown below, the speed limit decreases from 80kph to 60kph.

Figure 12
Ego_vehicle speed limit change scenario



Environmental requirements: A road that has at least one change in the speed limit.

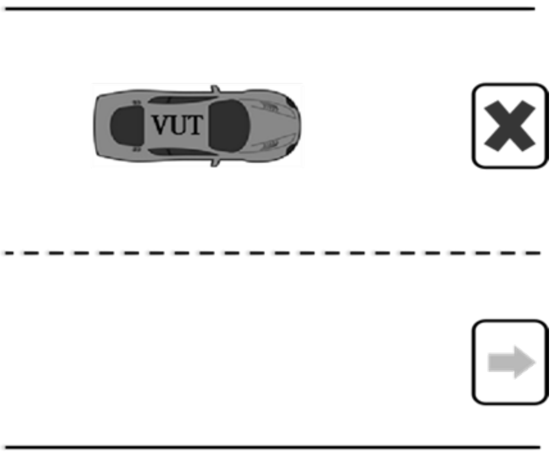
Ego vehicle behaviour: The ego vehicle drives on the road, presumably adapting its speed to the changing limitations.

Ego vehicle merge at lane end

(b) Signal lights

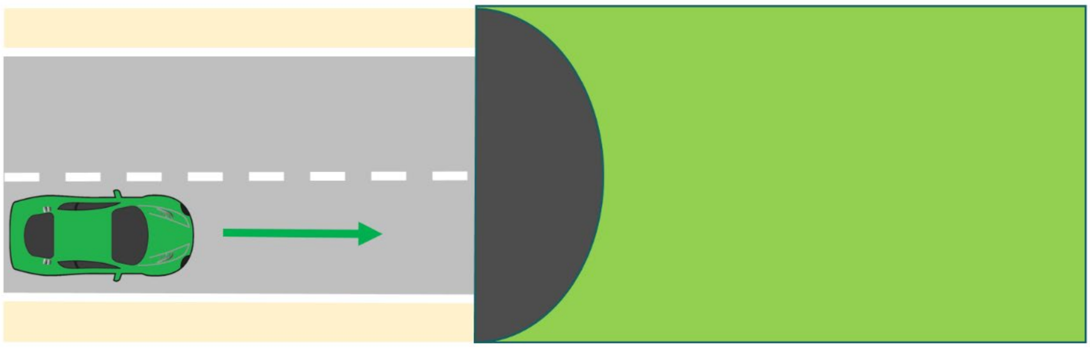
The test road consists of at least two lanes. The signal lights are set above the road, and the signal lights of adjacent lanes are kept in green state.

Figure 13
Testing scenario diagram for expressway signal lights



(c) Drive through tunnel

Figure 14
Schematic representation of driving through tunnel



General description:

The ego vehicle is driving through a tunnel (lack of GPS signals and natural light) [4]. The vehicle needs to adapt to the quickly changing light parameters and lack of global positioning. Depending on the speed of the ego vehicle, the difference between the light conditions outside and inside the tunnel and the length of the tunnel, the difficulty of the scenario is changing.

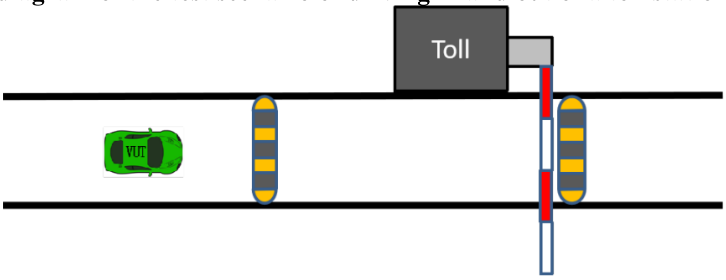
Emphasized scenario parameters: ego velocity, light conditions.

Tested parameters: ego lateral and longitudinal velocity, deviation to lane centre, etc.

(d) Toll

The test road is a long straight road with at least one lane. A toll station is set on this section, and toll station signs, speed limit signs and speed bumps are set in front of the toll station. This is shown in Figure 15.

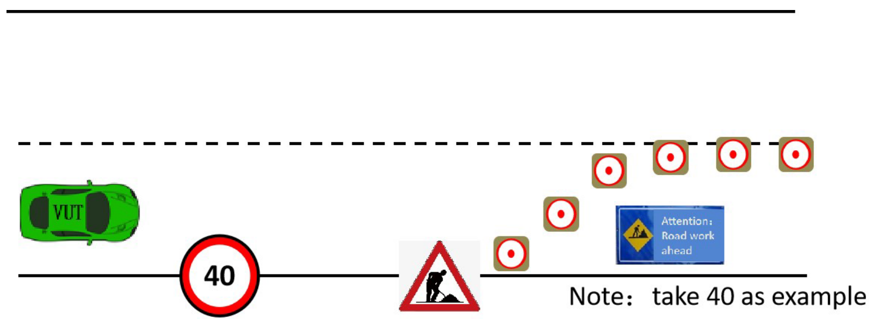
Figure 15
Schematic diagram of the test scenario of driving in and out of a toll station



(e) Conventional obstacles

The test road is a long straight road containing at least two lanes, and the middle lane line is a white dashed line. Within the lanes, conical traffic signs and traffic markings are placed according to the traffic control requirements of the road maintenance operation. This is shown in Figure 16.

Figure 16
Diagram of a conventional obstacle course.



D. Country specific road geometry

Note: This scenario is only applicable for limited countries or regions. Therefore, application of this scenario can be unnecessary depends on the target market of the ADS.

(a) Interceptor

For the ego vehicle, junctions present a challenge due to the increased likelihood of conflicts with other actors.

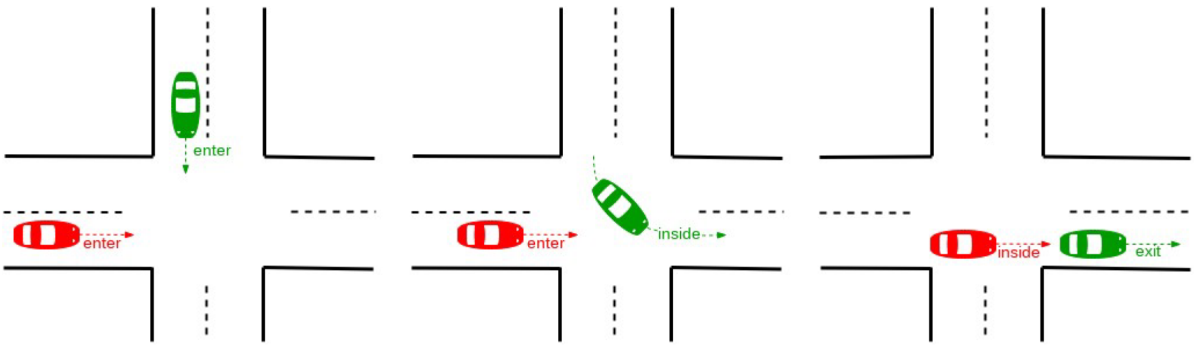
In this scenario, the ego vehicle traverses an intersection simultaneously with another car - the interceptor. This scenario tests the ego vehicle's behaviour when on a collision course with another car in an intersection, possibly with signs, signals, or traffic lights. The ego vehicle should be able to safely manoeuvre through the intersection and avoid or mitigate a collision.

Environmental requirements: A junction with at least three ways. It may or may not be controlled (i.e. have yield sign, traffic lights, etc.).

Ego vehicle behaviour: The ego vehicle traverses the junction in any direction (left, right or straight).

Other actors' behaviour: Another car approaches the same junction, from a different direction and traverses the junction such that its trajectory intersects with the ego vehicle's trajectory.

Figure 17
Interceptor scenario



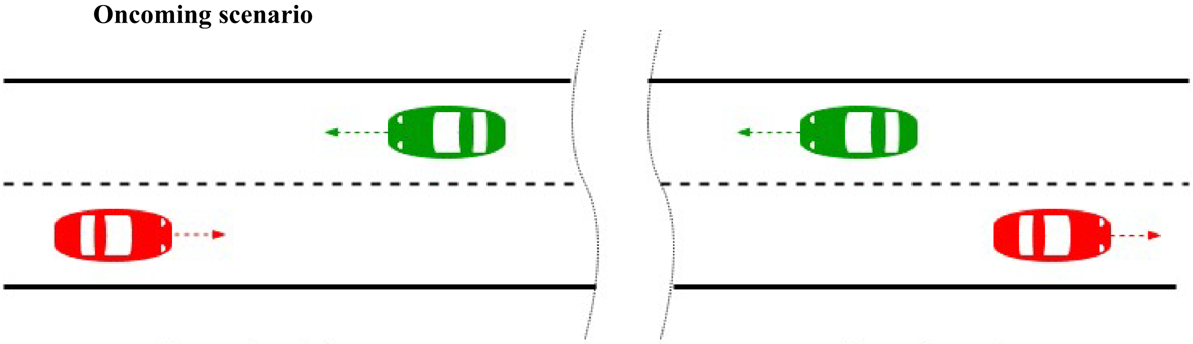
E. Unusual situation

Note: This scenario can happen in the real world. However, whether this kind of scenarios should be covered should be discussed in the appropriate group.

(a) **Wrong way driver (oncoming)**

Oncoming is a scenario in which a car approaches the ego vehicle from the opposite direction and drives past the ego vehicle.

Figure 18
Oncoming scenario



Environmental requirements: A two-lane road with traffic moving in opposite directions.

Ego vehicle behaviour: The ego vehicle drives in a lane, presumably at a constant speed.

Other actors' behaviour: At the start of the scenario, another car is in the opposing lane, approaching the ego vehicle. At the end, the other car is still in the opposing lane, having passed the ego vehicle.

VII. References

[1] ALKS Regulation UN R157. Available online at https://unece.org/transport/documents/2021/03/standards/un-regulation-no-157-automated-lane-keeping-systems-alks

[2] E. de Gelder, O. Op den Camp, N. de Boer, (The Netherlands): Scenario Categories for the Assessment of Automated Vehicles, Version 1.7, January 21, 2020.

[3] SAFE (Foretellix) Highway and ADAS Traffic Scenario Library, Scenario Definitions at the functional Level, Version 1.0, November 2020.

[4] Japan: Proposal of Traffic Scenarios for Highway Driving (Supplemental version for presentation), December 2020.

[5] China (CATARC): Proposal about functional scenario from catarc, December 2020.

[6] EC-JRC. Speed profile for car-following tests. Available online at https://wiki.unece.org/download/attachments/92013066/ACSF-25-13%20%28EC%29%2020190121_TestSpecification_ALKS_JRC.pdf?api=v2

[7] IGLAD 2019 Codebook, Conflict Types, 2019.

Annex III

Credibility assessment for using virtual toolchain in ADS validation

I. Introduction, motivation, and scope

1. The use of Modelling and Simulation (M&S) is becoming widespread thanks to the increasing computational capabilities, accuracy, usability, and availability of M&S software packages. M&S can be beneficial for ADS safety validation because it provides an opportunity to overcome some of the limitations of real testing and to increase the number of testing scenarios. Nonetheless, M&S can also lead to erroneous/seemingly correct results, especially in relation to complex simulations not adequately supported by robust practices addressing all M&S aspects beyond pure validation. Therefore, higher confidence in M&S credibility is needed so that virtual testing can be used instead of and in conjunction with the other pillars. In other words, M&S can be used for virtual testing if an assessor is able to consider the simulation results credible enough to make sound decisions taking into account the potential uncertainties of M&S.

2. If M&S is to be credible it needs to be validated. Validating the models and the simulation tools and process that make up M&S toolchain is difficult and there are limitations, which include the limited scope of the validation tests and the difficulty in gathering data to support the validation procedures. The use of M&S requires attention to all the factors influencing the quality and validity of M&S toolchain and all its separate components. The aim is to:

(a) Identify a common framework to determine, justify, assess and report the overall credibility of the M&S toolchain.

(b) Identify a way to indicate the levels of confidence in the results when a validation assessment takes place and also to determine the associated domains of applicability for the toolchain.

3. This framework should be general enough to be used for different M&S types and applications. Unfortunately, the goal is further complicated by the range and differences of ADS features and the variety of simulation tools and toolchains that are used. These considerations lead to the decision to use an (risk-based/informed) credibility assessment framework that can be applied to all M&S applications.

4. The proposed credibility assessment framework provides a general description of the main aspects needed for assessing the credibility of a M&S solution together with guidelines of the role played by the relevant assessor in the validation process with respect to credibility. The assessor should investigate the documentation and evidence supporting credibility during the audit phase. It is understood that the actual validation tests will take place once there is sufficient evidence that a simulation tool or toolchain produces credible results.

5. The outcome of the current credibility assessment will define the envelope in which the virtual tool can be used to support the ADS assessment.

II. Components of the credibility assessment framework

6. It is recommended that the M&S toolchain could be used for virtual testing if its credibility is established by evaluating its fitness for the intended purpose. It is recommended that credibility is achieved by investigating and assessing five M&S properties:

(a) Capability – what the M&S can do, and what are the associated risks;

(b) Accuracy – how well M&S does reproduce the target data;

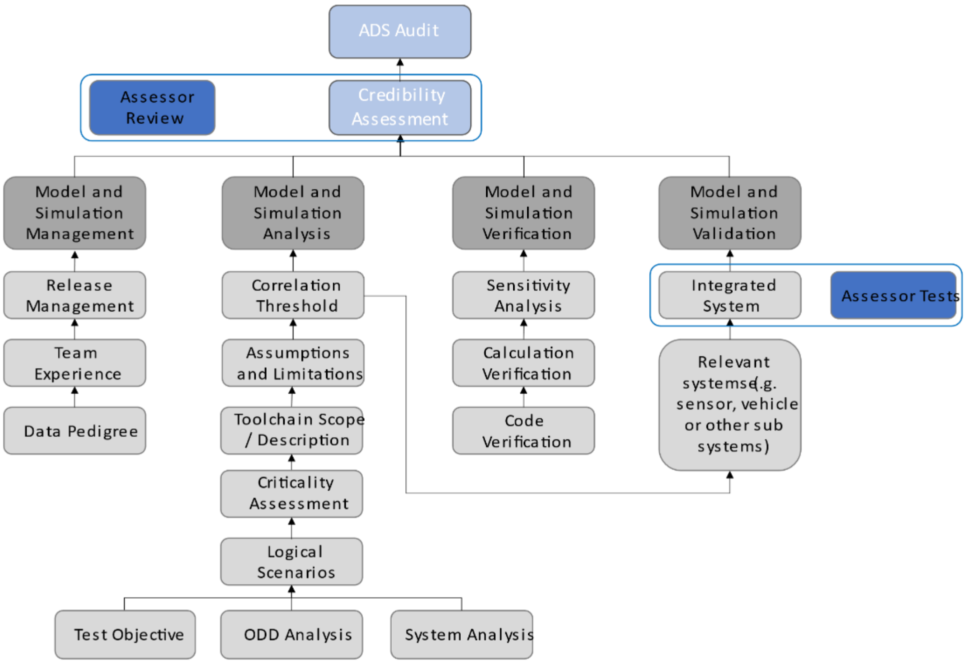
(c) Correctness – how sound & robust is the M&S data and the algorithm in the tools;

(d) Usability – what training and experience is needed and what is the quality of the process that manage its use.

(e) Fit for Purpose – how suitable is the M&S toolchain for the assessment of the ADS within its ODD.

Figure 1

Graphical representation of the relationships between the components of the credibility assessment framework



7. Therefore, credibility requires a unified method to investigate these properties and get confidence in the M&S results. The Credibility Assessment framework introduces a way to assess and report the credibility of M&S based on quality assurance criteria that allow an indication of the levels of confidence in results. In other words, the credibility is established by evaluating the key influencing factors that are the main contributors to the behaviour of the models and simulation tools and therefore affect the overall M&S toolchain credibility: The following all have an influence on the overall M&S credibility; organizational management of the M&S activity, team's experience and expertise, the analysis and description of the chosen M&S toolset, the pedigree of the data and inputs, verification, validation, uncertainty characterization. How well each of these factors is addressed indicates the level of quality achieved by M&S toolchain, and the comparison between the obtained levels and the required levels provides a qualitative measure of the M&S credibility and fitness for its use in virtual testing. A graphical representation of the relationship among the components of the credibility assessment framework is reported in Figure 1.

A. Models and Simulation Management.

8. The M&S lifecycle is a dynamic process with frequent releases that should be monitored and documented. As a result, it is recommended that management activities should be established to support the M&S in a work product management fashion. Relevant information on the following aspects should be included in this section.

9. It is recommended that this part should:

(a) Describe the modifications within the M&S toolchain releases,

(b) Designate the corresponding software (e.g., specific software product and version) and hardware arrangement (e.g., XiL configuration),

(c) Record the internal review processes that accepted the new releases,

(d) Be supported throughout the full duration of the virtual testing utilization

1. Releases management

10. It is recommended that any toolchain's version used to release data for certification purposes should be stored. The virtual models constituting the testing tool should be documented in terms of the corresponding validation methods and acceptance thresholds to support the overall credibility of the toolchain. The developer should establish and enforce a method to trace generated data to the corresponding toolchain version.

11. Quality check of virtual data. Data completeness, accuracy, and consistency are ensured throughout the releases and lifetime of an tool or toolchain to support the verification and validation procedures.

2. Team's Experience and Expertise.

12. Even though Experience and Expertise (E&E) are already covered in a general sense within an organization, it is important to establish the basis for confidence on the specific experience and expertise for M&S activities.

13. In fact, the credibility of M&S depends not only on the quality of the simulation models but also on the E&E of the personnel involved in the validation and usage of the M&S. For instance, a proper understanding of the limitations and validation domain will prevent possible misuse of the M&S or misinterpretation of its results.

14. In this perspective, important to establish the basis for the ADS manufacturer's confidence on the experience and expertise of:

(a) The teams that will validate the simulation toolchain and,

(b) The teams that will use the validated simulation for the execution of virtual testing with the purpose of validating the ADS.

15. Thus, if a team's E&E is good it increases the level of confidence and hence the credibility of M&S and its outcomes by ensuring that the human factors behind the M&S are taken into consideration and any possible human component risk is controlled, as expected, through its Management System.

16. If the ADS manufacturer's toolchain incorporates or relies upon inputs from organizations or products outside of the manufacturer's own team, it is recommended that the ADS manufacturer will include an explanation of measures it has taken to manage and develop confidence in the quality and integrity of those inputs.

17. The team's Experience and Expertise include two aspects:

(a) Organizational level:

18. The credibility is established by setting up processes and procedures to identify and maintain the skills, knowledge, and experience to perform M&S activities. The following processes should be established, maintained and documented:

(a) Process to identify and evaluate the individual's competence and skills;

(b) Process for training competent personnel to perform M&S-related duties

(b) Team level:

19. Once a toolchain has been finalized, its credibility is mainly dictated by the skills and knowledge of the individual/team that will validate the M&S and will use it for the validation of ADS. The credibility is established by documenting that these teams have received adequate training to fulfil their duties.

20. The ADS manufacturer should then:

(a) Provide the basis for the ADS manufacturer's confidence in the Experience and Expertise of the individual/team that validates the M&S toolchain.

(b) Provide the basis for the ADS manufacturer's confidence in the Experience and Expertise of the individual/team that uses the simulation to execute virtual testing with the purpose of validating the ADS.

21. The ADS manufacturer's demonstration of how it applies the principles of its Management Systems, e.g. ISO 9001 or a similar best practice or standard, with regard to the competence of its M&S organization and the individuals in that organization, will provide the necessary basis for this determination. It is recommended that the assessor not substitute its judgment for that of the ADS manufacturer with regard to the experience and expertise of the organization or its members.

4. Data/Input pedigree

22. The data/input pedigree contains a record of traceability about the ADS manufacturer's data used in the validation of the M&S.

(a) Description of the data used for the M&S validation

(a) The ADS manufacturer should document the data used to validate the models included in the tool or toolchain and note important quality characteristics;

(b) The ADS manufacturer should provide documentation showing that the data used to validate the models covers the intended functionalities that the toolchain aims at virtualizing;

(c) The ADS manufacturer should document the calibration procedures employed to fit the virtual models' parameters to the collected input data.

(b) Effect of the data quality (e.g. data coverage, signal to noise ratio, and sensors' uncertainty/bias/sampling rate) on model parameters uncertainty

23. The quality of the data used to develop the model will have an impact on model parameters' estimation and calibration. Uncertainty in model parameters will be another important aspect in the final uncertainty analysis.

5. Data/Output pedigree

24. The data/output pedigree contains a record of the M&S outputs used for the ADS validation.

(a) Description of the data generated by the M&S

(a) The ADS manufacturer should provide information on any data and scenarios used for virtual testing toolchain validation.

(b) The ADS manufacturer should document the exported data and note important quality characteristics e.g. using the correlation methodologies as defined Annex II.

(c) The ADS manufacturer should trace M&S outputs to the corresponding simulation setup:

(i) Effect of the data quality M&S credibility

(a) The M&S output data should be sufficient to ensure the correct execution of the validation computation. The data should sufficiently reflect the ODD relevant to the virtual assessment of the ADS.

(b) The output data should allow consistency/sanity check of the virtual models via possibly exploiting redundant information

(ii) Managing stochastic models

(a) Stochastic models should be characterized in terms of their variance

(b) The use of a stochastic models should not prohibit the possibility of deterministic re-execution

B. M&S Analysis and Description

25. The M&S analysis and description aim to define the whole toolchain and identify the parameter space that can be assessed via virtual testing. It defines the scope and limitations of the *models* and tools and the uncertainty sources that can affect its results.

1. General description

(a) ADS manufacturer should provide a description of the complete toolchain along with how the simulation data will be used to support the ADS validation strategy.

(b) The ADS manufacturer should provide a clear description of the test objective.

2. Assumptions, known limitations and uncertainty sources:

(a) The ADS manufacturer should motivate the modelling assumptions which guided the design of the M&S toolchain

(b) The ADS manufacturer should provide evidence on:

(i) How the manufacturer-defined assumptions play a role in defining the limitations of the toolchain;

(ii) The level of fidelity required for the simulation models.

(c) The ADS manufacturer should provide justification that the tolerance for simulation to real-world correlation is acceptable for the test objective

(d) Finally, this section should include information about the sources of uncertainty in the model. This will represent an important input to final uncertainty analysis, which will define how the model outputs can be affected by the different sources of uncertainty of the model used.

3. Scope (what is the model for?). It defines how the M&S is used in the ADS validation.

(a) The credibility of virtual tool should be enforced by a clearly defined scope for the utilization of the developed models.

(b) The matured M&S should allow a virtualization of the physical phenomena to a degree of accuracy which matches the fidelity level required for certification. Thus, the M&S will act as a “virtual proving ground” for ADS testing.

(c) Simulation models need dedicated scenarios and metrics for validation. The scenario selection used for validation should be sufficient such that there is confidence that the toolchain will perform in the same manner in scenarios that were not included in the validation scope.

(d) ADS manufacturers should provide a list of validation scenarios together with the corresponding parameter description limitations.

(e) ODD analysis is a crucial input to derive requirements, scope and effects that the M&S must consider to support ADS validation.

(f) Parameters generated for the scenarios will define extrinsic and intrinsic data for the toolchain and the simulation models.

4. Criticality assessment

26. The simulation models and the simulation tools used in the overall toolchain should be investigated in terms of their impact in case of a safety error in the final product. The proposed approach for criticality analysis is derived from ISO 26262, which requires qualification for some of the tools used in the development process. In order to derive how critical the simulated data is, the criticality assessment considers the following parameters:

(a) The consequences on human safety e.g. severity classes in ISO 26262.

(b) The degree in which the simulated results influence’s the ADS.

27. The table below provides an example criticality assessment matrix to demonstrate this analysis. ADS manufacturers may adjust this matrix to their particular use case.

Table 1
Criticality assessment matrix

Influence on ADS	Significant	N/A			
	Moderate				
	Minor				
	Negligible			N/A	
		Negligible	Minor	Moderate	Significant
		Decision consequence			

28. From the perspective of the criticality assessment, the three possible cases for assessment are:

(a) Those models or tools that are clear candidates for following a full credibility assessment;

(b) Those models or tools that may or may not be candidates for following the full credibility assessment at the discretion of the assessor;

(c) Those models or tools that are not required to follow the credibility assessment.

C. Verification

29. The verification of M&S deals with the analysis of the correct implementation of the conceptual/mathematical models that create and build up the overall toolchain. Verification contributes to the M&S's credibility via providing assurance that the individual tools will not exhibit unrealistic behaviour for a set of inputs which cannot be tested. The procedure is grounded in a multi-step approach described below, which includes code verification, calculation verification and sensitivity analysis.

1. Code verification

30. Code verification is concerned with the execution of testing that demonstrates that no numerical/logical flaws affect the virtual models.

(a) The ADS manufacturer should document the execution of proper code verification techniques, e.g. static/dynamic code verification, convergence analysis and comparison with exact solutions if applicable[3]

(b) The ADS manufacturer should provide documentation showing that the exploration in the domain of the input parameters was sufficiently wide to identify parameter combinations for which the M&S tools show unstable or unrealistic behaviour. Coverage metrics of parameters combinations may be used to demonstrate the required exploration of the model's behaviours.

(c) The ADS manufacturer should adopt sanity/consistency checking procedures whenever data allows

[3] Roy, C. J. (2005). Review of code and solution verification procedures for computational simulation. *Journal of Computational Physics*, *205*(1), 131-156.

2. Calculation verification

31. Calculation verification deals with the estimation of numerical errors affecting the M&S.

(a) The ADS manufacturer should document numerical error estimates (e.g. discretization error, rounding error, iterative procedures convergence);

(b) The numerical errors should be kept sufficiently bounded to not affect validation.

3. Sensitivity analysis

32. Sensitivity analysis aims at quantifying how model output values are affected by changes in the model input values and thus identifying the parameters having the greatest impact on the simulation model results. The sensitivity study also provides the opportunity to determine the extent to which the simulation model satisfies the validation thresholds when it is subjected to small variations of the parameters, thus it plays a fundamental role to support the credibility of the simulation results.

(a) The ADS manufacturer should provide supporting documentation demonstrating that the most critical parameters influencing the simulation output have been identified by means of sensitivity analysis techniques such as by perturbing the model's parameters;

(b) The ADS manufacturer should demonstrate that robust calibration procedures have been adopted and that this has identified and calibrated the most critical parameters leading to an increase in the credibility of the developed toolchain.

(c) Ultimately, the sensitivity analysis results will also help to define the inputs and parameters whose uncertainty characterization needs particular attention to characterize the uncertainty of the simulation results.

4. Validation

33. The quantitative process of determining the degree to which a model or a simulation is an accurate representation of the real world from the perspective of the intended uses of the M&S. Examples of virtual toolchain validation are reported in Annex 3 - Appendix 3. It is recommended that the following items be considered when assessing the validity of a model or simulation:

(a) Measures of Performance (metrics)

(a) The Measures of Performance are metrics that are used to compare the outputs of the virtual testing tool with real world performances. The Measures of Performance are defined during the M&S analysis.

(b) Metrics for validation may include:

(i) Discrete value analysis e.g. detection rate, firing rate;

(ii) Time evolution e.g. positions, speeds, acceleration;

(iii) Analysis of state changes e.g. distance/speed calculations, TTC calculation, brake initiation.

(b) Goodness of Fit measures

(a) The analytical frameworks used to compare real world and simulation metrics are generally derived as Key Performance Indicators (KPIs) indicating the statistical comparability between two sets of data.

(b) The validation should show that these KPIs are met.

(c) Validation methodology

(a) The ADS manufacturer should define the logical scenarios used for virtual testing toolchain validation. They should be able to cover to the maximum possible extent the ODD of virtual testing for ADS validation.

(b) The exact methodology depends on the structure and purpose of the toolchain. The validation may consist of one or more of the following:

(i) Validate Subsystem models e.g. environment model (road network, weather conditions, road user interaction), sensor models (Radio Detection And Ranging (RADAR), Light Detection And Ranging (LiDARs), Camera), vehicle model (steering, braking, powertrain);

(ii) Validate vehicle system (vehicle dynamics model together with the environment model);

(iii) Validate sensor system (sensor model together with the environment model);

(iv) Validate integrated system (sensor model + environment model with influences form vehicle model).

(d) Accuracy requirement

34. Requirement for the correlation threshold is defined during the M&S analysis. The validation should show that these KPIs are met. e.g. using the correlation methodologies as defined in Annex II.

(e) Validation scope (what part of the toolchain to be validated)

35. A toolchain consists of multiple tools, and each tool will use several *models*. The validation scope includes all tools and their relevant *models*.

(f) Internal validation results

(a) The documentation should not only provide evidence of the M&S validation but also should provide sufficient information related to the processes and products that demonstrate the overall credibility of the toolchain used.

(b) Documentation/results may be carried over from previous credibility assessments.

(g) Independent Validation of Results

36. The assessor should audit the documentation provided by the manufacturer and may carry out tests of the complete integrated tool. If the output of the virtual tests does not sufficiently replicate the output of physical tests, the assessor may request that the virtual and/or physical tests to be repeated. The outcome of the tests will be reviewed and any deviation in the results should be reviewed with the manufacturer. Sufficient explanation is required to justify why the test configuration caused deviation in results.

(h) Uncertainty characterisation

37. This section is concerned with characterizing the expected variability of the virtual toolchain results. The assessment should be made up of two phases. In a first phase the information collected from the "M&S Analysis and Description" section and the "Data/Input Pedigree" are used to characterise the uncertainty in the input data, in the model parameters and in the modelling structure. Then, by propagating all of the uncertainties through the virtual toolchain, the uncertainty of the model results is quantified. Depending on the uncertainty of the model results, proper safety margins will need to be introduced by the ADS manufacturer in the use of virtual testing as part of the ADS validation.

(i) Characterization of the uncertainty in the input data

38. The ADS manufacturer should demonstrate they have estimated the model's critical inputs by means of robust techniques such as providing multiple repetitions for the assessment of the quantity;

(ii) Characterization of the uncertainty in the model parameters (following calibration).

39. The ADS manufacturer should demonstrate that when a model's critical parameters cannot be fully determined they are characterized by means of a distribution and/or confidence intervals;

(iii) Characterization of the uncertainty in the M&S structure

40. The ADS manufacturer should provide evidence that the modelling assumptions are given a quantitative characterization by assessing the generated uncertainty (e.g. comparing the output of different modelling approaches whenever possible).);

(iv) Characterization of aleatory vs. epistemic uncertainty

41. The ADS manufacturer should aim to distinguish between the aleatory component of the uncertainty (which can only be estimated but not reduced) and the epistemic uncertainty deriving from the lack of knowledge in the virtualization of the process.

III. Documentation structure

42. This section will define how the aforementioned information will be collected and organized in the documentation provided by the ADS manufacturer to the relevant authority.

(a) The ADS manufacturer should produce a document (a "simulation handbook") structured using this outline to provide evidence for the topics presented;

(b) The documentation should be delivered together with the corresponding release of the toolchain and appropriate supporting data;

(c) The ADS manufacturer should provide clear reference that allows tracing the documentation to the corresponding parts of the toolchain and the data;

(d) The documentation should be maintained throughout the whole lifecycle of the toolchain utilization. The assessor may audit the ADS manufacturer through assessment of their documentation and/or by conducting physical tests.

Annex IV

List of occurrences recommended for reporting

43. Short term reporting is expected to be submitted for each critical occurrence. Periodic reporting is expected to be submitted in the form of aggregated data (per hour of operation or driven km) for ADS-vehicle type and related to ADS operation (i.e., when ADS is activated). If the manufacturer does not have access to complete operational information, it should agree how to proceed with the authority.

44. The following is a list of occurrences that have been derived from the ADS safety requirements set by FRAV. It is recommended that these form the basis of the reporting requirements. The occurrences have been subdivided into three categories, based on their relevance to the DDT, to the interaction with ADS vehicle users, and to ADS technical conditions. For each occurrence, its relevance to the short-term and/or periodic reporting has been flagged in the table below.

I. Occurrences related to ADS performance of the DDT, such as:

(a) Critical occurrences (as defined in the Glossary) known to the ADS manufacturer or OEM;

(b) Occurrences related to ADS operation outside its ODD;

(c) ADS failure to achieve a minimal risk condition when necessary;

(d) Communication-related occurrences (where connectivity is relevant to the ADS safety concept);

(e) Cybersecurity-related occurrences;

(f) Interaction with remote control centre (if applicable) related to major ADS or vehicle failures.

II. Occurrences related to ADS interaction with ADS vehicle users, such as:

(a) Driver unavailability (where applicable) and other user-related occurrences (e.g. user errors, misuse, misuse prevention);

(b) Occurrences related to Transfer of Control failure (reason, share compared to completed TOC);

(c) Prevention of takeover under unsafe conditions.

III. Occurrences related to ADS technical conditions, including maintenance and repair:

(a) Occurrences related ADS failure resulting in a request to intervene;

(b) Maintenance and repair problems;

(c) Occurrences related to unauthorized modifications (i.e. tampering);

(d) Modifications made by the ADS manufacturer or OEM to address an identified and significant ADS safety issue (with appropriate protections for related IP).

IV. Occurrences related to the identification of new safety-relevant scenarios

Table 1

Occurrences related to the identification of new safety-relevant scenarios

Occurrence	*Short-term reporting [1 Month]*	*Periodic Reporting [1 Year]*
1.a. Safety critical occurrences known to the ADS manufacturer or OEM	X	X
1.b. Occurrences related to ADS operation outside its ODD	X	X
1.c. ADS failure to achieve a minimal risk condition when necessary	X	X
1.d. Communication-related occurrences		X
1.e. Cybersecurity-related occurrences		X
1.f. Interaction with remote operator if applicable		X
2.a. Driver unavailability (where applicable) and other user-related occurrences		X
2.b. Occurrences related to Transfer of Control failure		X
2.c. Prevention of takeover under unsafe conditions		X
3.a. Occurrences related ADS failure		X
3.b. Maintenance and repair problems		X
3.c. Occurrences related to unauthorized modifications		X
3.d. Modifications made by the ADS manufacturer or OEM to address an identified and significant ADS safety issue	X (if the issue presented an unreasonable risk to safety)	X
4. Occurrences related to the identification of new safety-relevant scenarios	(already covered under 1.a, 1.b, 1.c and 3,d)	X

Annex V

Outline of the approach for the testing methods for track testing and real world testing

I. Introduction

1. An initial overview of best practices, procedures, technical resources and tools related to track testing and real world testing was prepared.[4]

2. The overview showed that numerous test procedures and standards for track testing have been developed and used to assess the safety of vehicles with automated driving systems (e.g. ALKS) and particularly with advanced driver assistance systems, which can serve as input to the to-be-developed track testing methodology.

3. The overview furthermore showed that no test procedure to assess the safety of vehicles with automated driving systems on public roads has been developed yet[5], with most of the available documentation concerning guidance or specifications on testing (i.e. trials) such vehicles by OEMs during the developmental stages of their systems, or the testing of human drivers.

4. This annex outlines the suggested approach for track and real world testing : test matrices, Consideration and recommendations on the next steps for the development of the testing methods are provided in Annex VI.

II. The test matrices

5. The starting point for the development of the methods for track testing and real world testing is the test matrix approach. This approach recommends the use of one general matrix for physical testing, as well as two test matrices specifically designed for respectively track testing and real world testing.

6. The purpose of the general matrix for physical testing would be to provide a clear overview of how the respective safety performance requirements set by FRAV could be assessed using track testing, real world testing, or both.[6]

7. The test matrices for respectively track testing and real world testing would differ in design, in order to take into account the different settings in which the tests are conducted, as well as to ensure that the strengths of each testing method can be utilized.

Please note that the example test matrices set out in this annex are merely illustrative and therefore include mock-up criteria.

A. The general matrix for physical testing

8. The general matrix would provide a clear overview of the type or types of physical testing to be used for assessing compliance with the applicable safety requirements set by FRAV.

---

[4] Working paper VMAD-SG4-06-05

[5] UN Regulation No. 157 on automated lane keeping systems (ALKS) provides provisions for a real world test. For the purpose of developing the NATM's real world test, these provisions are however not detailed enough to be regarded as specifications including a procedure.

[6] This general matrix for physical testing would only include applicable safety performance requirements suitable for physical testing, excluding those that are only to be assessed using other pillars of the NATM. Should VMAD decide to establish a general overview at VMAD/NATM level setting out which pillar/pillars should be used to assess compliance with the respective safety performance requirements set by FRAV, then this general matrix for physical testing could be integrated into such overview.

9. The example in table VII.1 illustrates the basic concept for the overview based on (a selection of) the initial 40 safety topics drafted by FRAV.[7] Please note that the example is merely illustrative and should not be regarded as VMAD's position on the applicability of each test method per safety topic.

10. Moreover, the example does not take into account any further development on the safety topics by FRAV since the 18th VMAD meeting. The safety topics are furthermore expected to be set out in more detail in the future, each topic containing one of more measurable requirements.

11. It is recommended that, once developed, these measurable requirements would be listed in the left column of the table instead of the currently listed safety topics.

Table 1
Example of the General Matrix for Physical Testing

(FRAV) Safety Requirement	*Track Testing*	*Real World Testing*
1. The ADS should perform the entire Dynamic Driving Task.	Yes	Yes
2. The ADS should control the longitudinal and lateral motion of the vehicle.	Yes	Yes
(…)		
7. The ADS should adapt its behavior in line with safety risks.	Yes	If encountered
8. The ADS should adapt its behavior to the surrounding traffic conditions.		Yes
(…)		
30. The ADS should safely manage short-duration ODD exits.	Yes	Yes
31. Pursuant to a collision, the ADS should stop the vehicle and deactivate.	Yes	If encountered
(…)		

12. 'If encountered', as used in the table above, would indicate that real world testing would not seek to assess the particular requirement but would do so if it occurred. Some situations are clearly undesirable from a safety perspective on public roads. However, given that random traffic situations are encountered during real world testing, such traffic situations could organically occur and in this case, the performance with regards to the specific requirement should be assessed. Safety during testing on public roads should also be taken into account, and the assessor or the driver should ensure they can take over the driving task if needed.

13. VMAD's SG4 will elaborate in a future activity, how the assessment of such 'if encountered' occurrences should be integrated in the testing method for real world testing, e.g. whether they would be included in the test matrix or whether the testing protocols would provide guidance/instructions on how assessors are expected to handle such cases.

14. Instead of "Yes" or "If encountered", the table could also be structured to provide more information on the intended purpose/aim of the test. For example:

Table 2
Example of another structure for the general matrix for physical testing

[7] As set out in working paper VMAD-18-03.

(FRAV) Safety Requirement	*Track Testing*	*Real World Testing*
XX. The ADS should respond safely to the cut-in of another vehicle.	Verification of the ADS crash-avoidance response to a dangerous cut in.	Nominal verification that the ADS adapts the vehicle positioning in response to the cut in. Verification of the ADS crash-avoidance response to a dangerous cut in, if encountered.

B. The test matrix for track testing

15. The left column of the test matrix for track testing will refer to scenarios developed by VMAD's SG1.VMAD's SG4 anticipates these would include the traffic situation, infrastructure elements, objects, ODD elements, etc.

16. The safety requirement(s) column will cross-reference the applicable safety requirement(s), which will be set out by FRAV, and will be assessed in the respective scenario. VMAD's SG4 anticipates that FRAV will provide requirements enabling determinations of the pass/fail criteria, which will in turn be set out in the assessment specification column.

17. If applicable, the additional test specification column will allow for any additional conditions or parameters to be specified, which were/could not be described in either the traffic scenario or the safety requirement(s), but are necessary to conduct the track test (e.g. minimum duration of the test).

18. Please note that the example matrix on the next page is only illustrative of the envisaged structure. The content is intentionally non-specific and should not be regarded as VMAD's position on the suitability of using track testing to assess the listed safety requirements.

19. The eventual scenarios, safety requirement(s) and assessment specifications are to be sourced from VMAD's SG1 and FRAV, with any additional test specifications to be added based on discussions within VMAD's SG4.

Table 3
Example of a test matrix for track test

Traffic Scenario	*Safety Requirement(s)*	*Additional Test Specifications*	*Assessment Specification*
This column would cross-reference the testing with the scenario upon which the testing is based. VMAD's SG4 anticipates that the scenarios would cover the traffic situation, infrastructure elements, objects, ODD elements, etc.	This column would cross-reference the testing with the safety requirements relevant to the traffic scenario. SG4 anticipates that FRAV would provide requirements enabling determinations of the pass/fail criteria, to be set out in the assessment specification column.	This column would complement the description of the traffic scenario with additional information or parameters necessary for conducting the track test, if applicable.	This column would set out the assessment specification.
The following examples illustrate the concept of the matrix for track testing. VMAD's SG4 has intentionally provided non-specific examples. The scenarios and safety requirements would be sourced from VMAD's SG1 and FRAV. The matrix would evolve in line with progress of these activities.			
Unobstructed travel on a straight path	• Safe lateral positioning in a lane of travel	• A minimum test duration of 5 minutes	The test shall demonstrate that the ADS does not leave its lane and maintains a stable position inside its ego lane across the speed range within its system boundaries.
Unobstructed travel along a curve	• Safe lateral positioning in a lane of travel • Adapt to road conditions	• A minimum test duration of 5 minutes	The test shall demonstrate that the ADS does not leave its lane and maintains a stable position inside its ego lane across the speed range and different curvatures within its system boundaries.
Cut-in by another vehicle while traveling on a straight path	• Respond safely to the cut-in • Safe longitudinal positioning relative to a lead vehicle	• Scenario with selected parameters to verify the ADS crash-avoidance response to a dangerous cut in per the safety requirements[8]	The test shall demonstrate that the ADS is capable of avoiding a collision with a vehicle cutting into the lane of the ADS vehicle up to a certain criticality of the cut-in manoeuvre.
ODD exit scenario	• ADS detection of ODD boundary • Automated response (if failed fallback user response or no fallback user)	• Test for failed fallback user response	The test shall demonstrate that the ADS is capable of bringing the vehicle to a safe stop, in case of a failed fallback user response.

[8] This inclusion assumes the traffic scenario does not prescribe the (range of parameters to be selected for the) occurrence of a safety-critical situation. If that were to be included in the scenario, this field could be empty.

C. The test matrix for real world testing

20. The left columns of the test matrix for real world testing would set out the measureable safety requirements to be developed by FRAV.

21. The top rows on the right side would set out the traffic situations required to be encountered during real world testing. Given the dynamic nature of traffic on public roads, it is considered unlikely that traffic situations will occur exactly as described in the traffic scenarios developed by VMAD's SG1, and therefore they are not referenced in the test matrix. Instead, the traffic situations listed in the second row will be described in the testing protocols accompanying the test matrix. These envisaged description will be rather general in order to ensure that there is a very high probability of them being encountered during real world testing. In order to prevent confusion with the term "traffic scenarios" developed by VMAD's SG1, they have (provisionally) been termed as "traffic situations" instead.[9]

22. Please note that the five traffic situations set out in the example are merely illustrative.

23. The remaining fields of the matrix represent the assessment specification per safety requirement for the applicable traffic situations, which are to be sourced from FRAV. The assessment specification would summarize the desired performance in one sentence, with a more detailed description of the assessment specification to be set out in the testing protocols accompanying the test matrix, where necessary.

24. The inclusion of an assessment specification reflects the need for each safety requirement to be verified for its respective traffic situation. As an illustration, in the example table's Row 1.1, the compliance with the safety requirement on lane keeping would have to be verified in all the traffic situations. (Please note that the assessment specifications in the table are merely illustrative and moreover do not reflect VMAD's position on whether compliance with the safety requirement should be verified in the respective traffic situations for which the mock-up assessment specifications are provided).

25. As previously mentioned, VMAD's SG4 will continue to discuss how the "If encountered" situations can be appropriately reflected in the test matrix. This may occur in two situations. First, the assessment of safety requirements that are undesirable to be conducted on public roads, but which may nevertheless occur.[10] Second, the assessment of safety requirements (during nominal traffic conditions) that cannot be assured (and therefore required) to be encountered during real world testing, but which may occur.

26. An illustration of the first is the example on Row 2.1 of the table on the safe response to a cut-in. The requirement is the assessment of the ADS' response to a (nominal) cut-in of another vehicle during real world testing. The ADS' response to a dangerous cut-in could only be assessed if encountered during real world testing, as signalled by the addition of ', if applicable.'.

27. An illustration of the second is also the example in Row 2.1 of the table on the safe response to a (nominal) cut-in. This situation is likely but not guaranteed to occur in any or possibly all of the traffic situations listed in the top row of the table. When it does occur it should be assessed.

28. For both cases, VMAD's SG4 will further discuss what the most efficient and clear way would be to signal such 'Assess if encountered' requirements. Suggestions made so far include:

(a) Handle such occurrences separately from the test matrix (e.g. only provide guidance/instructions in the testing protocols);

---

[9] Should VMAD's SG1 develop general scenarios suitable for use in the test matrix for real world testing, VMAD's SG4 would consider references to those general scenarios instead.

[10] It should be possible for the assessor to interrupt the test on public roads, should the situation become dangerous. VMAD's SG4 will further discuss this topic and may decide to provide guidance in the testing protocols.

(b) Include them in the test matrix itself, but signal their conditionality (e.g. Row 2.2 in the example table. Note in particular the assessment specification for lane changes, where both a required assessment and a conditional 'if encountered' assessment are included);

(c) Signal the existance of an 'If encountered' assessments (e.g. using '*'), but setting out the conditional assessment specification as well as guidance/instructions in the testing protocols. (Please note that this latter option has not been illustrated in the example table).

29. Aspects related to routing (e.g. minimum duration, minimum frequency of a given traffic situation encountered during testing, etc.) would be set out in the accompanying test protocols.

Table 4

Example of a test matrix for real world testing: motorway application

		Traffic Situations				
Safety Requirements		*Driving on the motorway*	*Merging*	*Lane Change*	*Overtaking*	*Exiting Motorway*
1.1.	Safe lateral positioning in a lane of travel	The ADS demonstrates it does not leave its lane and maintains a stable position inside its ego lane across the speed range within its system boundaries.	The ADS demonstrates it achieves a stable position inside the target lane upon completion of the lane change procedure.	The ADS demonstrates it achieves a stable position inside the target lane upon completion of the lane change procedure.	The ADS demonstrates it achieves a stable position inside the target lane upon completion of the lane change procedure.	The ADS demonstrates it maintains a stable position in the off-ramp lane.
2.1.	Respond safely to the cut-in of another vehicle	The ADS adapts the vehicle positioning in response to the (nominal) cut in. The ADS responds appropriately[11] to a dangerous cut in, if applicable.[12]				
2.2.	Safe longitudinal positioning relative to a lead vehicle	The ADS demonstrates it maintains a safe longitudinal position relative to a lead vehicle.	The ADS demonstrates it maintains a safe longitudinal position relative to a lead vehicle during and upon the completion of the lane change procedure.	The ADS demonstrates it maintains a safe longitudinal position relative to a lead vehicle prior and during the lane change procedure. The ADS demonstrates it maintains a safe longitudinal position relative to a lead vehicle upon the completion of the lane change procedure, if applicable.	The ADS demonstrates it maintains a safe longitudinal position relative to a lead vehicle prior and during the lane change procedure.	The ADS demonstrates it maintains a safe longitudinal position relative to a lead vehicle, if applicable.

[11] What constitutes an 'appropriate response' would then be set out in the testing protocols that accompany the test matrix, sourced from FRAV.

[12] To be determined whether 'If encountered' situations should be included in the matrix itself. Included here, as well as in other parts of the table, as an illustration.

Annex VI

Areas for future work

Scenarios catalogue — updating the catalogue

1. The text below discusses future work that may be conducted to develop and maintain a VMAD scenarios catalogue:

2. If scenarios not covered by the scenario catalogue are identified and deemed necessary, they should be included in the scenario catalogue.

3. It is envisaged that a scenario catalogue will have tags for all scenarios corresponding to their relevant ODD attributes (using a standardised ODD taxonomy) and behaviour competencies.

4. Country specific scenarios should be respected and need to be covered in the scenario catalogue in the long term.

5. The scenario catalogue is not necessarily exhaustive and authorities may need to consider additional scenarios as necessary to support safety validation of an ADS feature. [Such a decision could be based on the ODD and behaviour competencies of the ADS. For example, if an AV is developed with an ODD which is not covered in the scenario catalogue, it is essential to add new scenarios to the catalogue to ensure the scenarios used for testing are a function of the ODD.]

Pillars 2 and 3 — track and real world testing — considerations and next steps

6. The next step in the development of the test methods for track and real world testing (described in Annex V) is populating the test matrices with safety requirements, traffic scenarios/traffic situations and the assessment specifications. However, it will merely be the first of several steps before the test matrix approach can be used as an assessment method.

7. This section therefore outlines the next steps that are required in order to operationalize the test matrix approach, together with some initial considerations.

A. Populating the test matrix

8. In order to be able to advance with the development process of the test matrix testing method, it is first necessary to populate the test matrices with requirements, scenarios and assessment specifications. This is because most, if not all of the subsequent steps largely depend on the content of the matrix itself. For example, without knowing what will be required to be tested and against which criteria, it would difficult, if not impossible, to determine the length and scope of the real world testing aspect.

9. The test matrix would be populated with the requirements and assessment specifications to be developed by FRAV, and for track testing the scenarios developed by VMAD's SG1 as well. Given that FRAV and VMAD's SG1 are currently still in the process of developing respectively the requirements and traffic scenarios, SG4's work on the matrixes themselves will be largely on hold until the requirements and scenarios become available.

10. With regards to the populating the test matrices in due time, the criteria to be included for testing would be selected in coordination with VMAD and FRAV, whereas the scenarios would be selected in coordination with SG1.

B. Developing the test protocols

11. Once a test matrix has been populated, the accompanying test protocols[13] will be developed by VMAD's SG4. These test protocols would include, for example, the scope and length of testing, conditions for testing and routing (as far as not provided for by either the criteria or traffic scenario/traffic situation descriptions). It would also include other aspects that are necessary for the persons conducting the testing to ensure a consistent interpretation of the test matrix and protocols and to deliver harmonized assessments.

C. Validation of the testing approach

12. The test matrices and accompanying test protocols first need to be validated during a pilot or trial phase. This would ensure that the approach is providing the desired assessment of the ADS vehicles. These validations are particularly important for real world testing, as no regulatory framework, procedure, or specification currently exists to assess the safety of the ADS.

13. This validation process will be developed further in a later stage, once (suitable drafts of) the test matrices and accompanying test protocols have been developed. There are a number of questions that need to be considered during the development of the validation process including:

(a) How many test organizations and test vehicles are required?

(b) How many times would the trials need to be repeated?

(c) Who conducts these trials?

(d) How many countries need to validate the test matrix and test protocols?

(e) Would each country need to conduct their own trials?

Pillar 5 — In-service monitoring and reporting — reporting from other sources

14. The effectiveness of the ISMR pillar will be determined by the availability of data on ADS safety performance. This means that limiting the reporting requirements to manufacturers only would also limit the type of occurrences that can be covered by ISMR, and consequently the level of safety improvement achievable through operational experience feedback. Other transport sectors extend the operational reporting mechanism to other sources including drivers, operators, users, managers, and any other person connected to the vehicle's operation. Discussion on the possibility of similar extensions requires exchanges between WP.29 and WP.1.

15. For example, occurrences related to traffic rules infringement cannot be covered with only data collected from the vehicle as the expectation is that the ADS will not intentionally infringe the law. Therefore, if it is not aware that it has infringed the law it will not record any data. It may therefore be necessary to consult with other parties, such as local authorities and ADS vehicle's user(s), as well as the manufacturers to identify and gather data on this category of occurrences.

Pillar 5 — In-service monitoring and reporting — information sharing among safety authorities/contracting parties

16. The final aim of the ISMR pillar is to improve ADS safety through dissemination of lessons learned in the form of safety recommendations. If this information is shared promptly

[13] Test parameters should take into account the ODD of the ADS under test.

and widely then it should have an impact at both a national and international level. Safety authorities should have access to such recommendations as well as the reporting from manufacturers, plus other relevant information (e.g., data from highway authorities, crash investigations, research, national statistics, etc.). This should allow them to react to any safety issues associated with ADS deployment.

17. A mechanism to share information across safety authorities at a global level is desirable and could be coordinated by GRVA/VMAD, under WP.29 directions.