# Outcome of the GRVA workshops on Artificial Intelligence and Vehicle Regulations

## I.  Mandate

1.  Following the AC.2 decisions of November 2020 and the discussions at the last sessions of GRVA, GRVA requested the secretariat to organize a technical workshop focusing primarily on definitions for Artificial Intelligence, relevant for GRVA activities. The first workshop took place on 18 March 2022. The experts agreed to convene a second workshop on 9 May 2022 to explore the AI use cases and their relevance for GRVA with regards to safety.

2.  The experts discussed whether technology neutral performance requirements are sufficient for the purpose of GRVA or if specific provisions would be necessary. The experts developed draft definitions, drafted a table with use cases and their relevance with regards to vehicle regulations and reflected on the potential activities that could be necessary in the framework of the New Assessment Test Method developed by GRVA and its IWG on Validation Method for Automated Driving (VMAD).

## II.  List of AI relevant definitions in the context of vehicle regulations

3.      The terms below are inspired by the definitions under review at the International Standard Organization (see ISO/IEC 22989).

[4.      **Artificial intelligence** is a set of methods or automated entities that together build, optimize and apply a model so that the system can, for a given set of predefined tasks, compute predictions, recommendations, or decisions.

5.      **Machine learning** is a data based computational techniques to create an ability to "learn" without an explicitly programmed algorithm such that the model's behaviour reflects the data or experience.

6.      **Machine learning model** is a mathematical construct that generates an inference, or prediction, based on input data.

7.      **Deep learning** is an approach to creating rich hierarchical representations through the training of neural networks with many hidden layers.

8.      **Supervised learning** is a type of machine learning that makes use of labelled data during training.

9.      **Unsupervised learning** is a type of machine learning that makes use of unlabelled data during training.

10.      **Reinforcement learning** is a type of machine learning utilizing a reward function to optimize a machine learning model by sequential interaction with an environment.

11.      **Dataset** is a collection of data with a shared format and goal-relevant content.

12.      **Data sampling** is a process to select a subset of data samples intended to present patterns and trends similar to that of the larger dataset being analysed.

13.      **Data annotation** is the process of attaching a set of descriptive information to data without any change to that data.

14.      **Training** is the process to establish or to improve the parameters of a machine learning model, based on a machine learning algorithm, by using training data.

15.     **Retraining** is an approach to creating rich hierarchical representations through the training of neural networks with many hidden layers.

16.     **Continuous learning** describes incremental training of an AI system throughout the lifecycle to achieve defined goals governed by pre and post operation risk acceptance criteria and human oversight.

17.     **Self-learning** describes incremental training of an AI system throughout the lifecycle to achieve defined goals governed by pre and post operation risk acceptance criteria making possible a continuous activation of the new system output with or without human oversight.

18.     **Online learning** describes incremental training of a new version of the AI system during operation to achieve defined goals based on post operation acceptance criteria and human oversight without activating the new system output until released.

19.     **Human oversight** is AI system property guaranteeing that built-in operational constraints cannot be overridden by the system itself and is responsive to the human operator, and that the natural persons to whom human oversight is assigned.

20.     **AI lifecycle** consists out of the design and development phase of the AI system, including but not limited to the collection, selection and processing of data and the choice of the model, the validation phase, the deployment phase and the monitoring phase. The life cycle ends when the AI system is no longer operational.

21.     **Safe-by-design** is system property enabled by development and lifecycle activities to claim system measures bring risks to an acceptable level.

22.     **Trustworthiness** is the ability to meet stakeholders' expectations in a verifiable way.

23.     **Bias** is a systematic difference in treatment of certain objects, people, or groups in comparison to others.

24.     **Fairness / Fairness matrix** is a way of describing bias.

25.     **Predictability** is a property of an AI system that enables reliable assumptions by stakeholders about the output.

26.     **Reliability** is a property of consistent intended behaviour and results.

27.     **Resilience** is the ability of a system to recover operational condition quickly following an incident.

28.     **Robustness** is the ability of a system to maintain its level of performance under any circumstances.

29.     **Transparency of an organization** is the property of an organization that appropriate activities and decisions are communicated to relevant stakeholders in a comprehensive, accessible and understandable manner.

30.     **Transparency of a system** is property of a system to communicate information to stakeholders.

31.     **Explainable** means a property of an AI system to express important factors influencing the AI system that results in a way that humans can understand.

32.     **Black/Grey/White box [testing]** are [tests of] systems / software in which functionality are unknown / partially know / known.]

# III. AI use cases in the automotive sector

Note: The following table was prepared by the experts from CLEPA and OICA

An editable version of this table is available here: https://unece.org/transport/events/grva-technical-workshop-artificial-intelligence-2nd

| | AI Application | | Non Safety functions<br><br>e.g. Infotainment<br>Out of Scope of type approval | Safety functions | | | Non Driving Functions |
|---|---|---|---|---|---|---|---|
| | | | | Driving Function | | | |
| | | | | Perception | Planning | Actuation | |
| Conventional Software | Artifical Intelligence (AI) | Artificial Intelligence is a set of methods or automated entities that together build, optimize and apply a model so that the system can, for a given set of predefined tasks, compute predictions, recommendations, or decisions | Natural language processing | Out of Scope [Non-AI]<br><br>Detection of other road users for AEBS, ACC<br>Detection of road infrastructure for LDW, LKAS | Out of Scope<br><br>Activation of FCW and AEBS based on ego vehicle position and other road users | Not Applicable | Out of Scope<br><br>Detection of driver's face for ID (under conditions ensuring privacy) |
| Artificial Intelligence | Supervised Learning (SL) | Supervised learning is a type of machine learning that makes use of labelled data during training | Gesture control<br>Voice Recognition | Detection of other road users for AEBS, ACC<br>Detection of passive road infrastructure for LDW, LKAS | Trajectory prediction using drivable path prediction from labelled data (e.g. HD maps) | Not Applicable | Detection of drivers eye gaze / state for DMS<br>Fault detection, Predictive Maintenance |
| | Unsupervised Learning (UL) | Unsupervised learning is a type of machine learning that makes use of unlabelled data during training | | Streamlining data labelling process for less safety critical systems like ISA.<br>Extracting scenarios from real world data to suport validation<br>Generation of synthetic data for supervised learning / distortion of real world data | Trajectory prediction using Kalman filters, KalmanNet or Gaussian Process architectures, or other architectures | Not Applicable | [?] |
| | Semi Supervised Learning (SSL) | Semi supervised learning is a technique that "learns" from a mix of labelled data and data that is both un-labelled and unstructured. They build on a small set of known exemplars and then use this information to guide unsupervised learning. | | Streamlining data labelling process for less safety critical systems like ISA. | Shadow mode' used in development for training control algorithms | Not Applicable | [?] |
| | Reinforcement Learning (RL) | Reinforcement learning is a type of machine learning utilizing a reward function to optimize a machine learning model by sequential interaction with an environment | | Some manufacturers are starting to use RL for perception, could potentionally be used in cooperative perception in the future. | Lane Centering or ACC systems may use RL due to the reduction in cost / data required to train the system | Not Applicable | Predictive Maintenance |

# IV. Impact of Artificial intelligence on the New Assessment Test Method

Note: The following table was prepared by the experts from CLEPA and OICA.

An editable version of this table is available here: https://unece.org/transport/events/grva-technical-workshop-artificial-intelligence-2nd