Industry input on GRVA-13-04 titled “Outcome of the GRVA workshops on Artificial Intelligence and Vehicle Regulations”

The input is marked in “tracked change”.

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. Relevance for GRVA

3. This short chapter provides two examples aimed at suggesting that GRVA might have to look into Artificial Intelligence in the context of vehicle regulations.

A. Test results reproducibility according to UN GTRs ad UN Regulations

4. GRVA develops technical requirements and guidance that are technology neutral, unless a specific technology requires appropriate and specific provisions.

5. GRVA discussed (GRVA-12-06) that in the case of functions, which are based on software that is generated by Artificial Intelligence, the outcome associated with this AI for a given situation will not necessarily be predictable.

6. The predictability of test results is an important factor for the type-approval and for the self-certification.

B. Specific features of AI systems used in automotive products

7. AI systems, used in automotive products, may provide the possibility for offline retraining combined with a thorough validation and Over-the-Air (OTA) updates. This offers a compromise that allows adaptations to model drift and model staleness processes while guaranteeing a certain level of safety and security.

8. GRVA might wish to evaluate whether the provisions regarding software updates (in UN Regulation No. 156 and in the recommendations on uniform provisions concerning cyber security and software updates) adequately address retraining and OTA updates.
III. List of AI relevant definitions in the context of vehicle regulations

9. The terms below are are taken from the definitions under review at the International Standard Organization (see ISO/IEC 22989).

10. 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.

11. Machine learning is a collection of data-based computational techniques to create an ability to learn without an explicitly programmed outcome such that the model's behavior reflects the data or experience.

12. Machine learning model is a computer science construct that generates an inference, or prediction, based on input data.

13. Deep learning is a process whereby neural networks use single/multiple layers of processing intended to extract progressively higher level features from data.

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

15. Unsupervised learning is a type of machine learning that makes use of unlabeled data during training.

16. Reinforcement learning is a (collection of) training technique that permits to an agent to learn actions to be taken from experiences, optimizing a quantitative reward gained along the time.

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

18. Data sampling is a process to select a subset of data intended to present patterns and trends similar to those of the larger dataset being analyzed.

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

20. Training is the process to tune the parameters of a machine-learning model.

21. Training data is a subset of input data samples used to train a machine learning model

22. Validation data is data used to assess the performance of a final machine learning model

23. 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.

24. 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.

25. 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.

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

27. Trustworthiness is the ability to meet stakeholders’ expectations in a verifiable way.
28. **Bias** is a systematic difference in treatment (including categorization/observation) of certain objects, people, or groups in comparison to others.

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

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

31. **Reliability** is a property of consistent intended behavior and results.

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

33. **Robustness** is the ability of a system to maintain its level of performance during the whole lifecycle.

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

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

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

37. **Black box** is a systems / software in which the detailed functionality is unknown.

38. **Grey box** is a systems / software in which the detailed functionality is partially known.

39. **White box** is a systems / software in which the detailed functionality is known.

40. **Black/Grey/White box testing** are tests of systems / software in which functionalities 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

<table>
<thead>
<tr>
<th>AI Application</th>
<th>Non Safety functions</th>
<th>Safety functions</th>
<th>Non Driving Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>e.g. Infotainment</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Out of Scope</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>of type approval</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Artificial Intelligence (AI)</strong></td>
<td>Natural language</td>
<td>Out of Scope</td>
<td>Out of Scope</td>
</tr>
<tr>
<td></td>
<td>processing</td>
<td>[Non-AI]</td>
<td>(under conditions ensuring privacy)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Supervised Learning (SL)</strong></td>
<td>Gesture control</td>
<td>Detection of other road users for AEBS, ACC</td>
<td>Trajectory prediction using drivable path prediction from labelled data (e.g. HD maps)</td>
</tr>
<tr>
<td></td>
<td>Voice Recognition</td>
<td>Detection of road infrastructure for LDW, LKAS</td>
<td>Not Applicable</td>
</tr>
<tr>
<td><strong>Unsupervised Learning (UL)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Semi Supervised Learning (SSL)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Reinforcement Learning (RL)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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.

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

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

Semi supervised learning is a technique that "learns" from a mix of labelled data and data that is both unlabelled and unstructured. They build on a small set of known exemplars and then use this information to guide unsupervised learning.

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 potentially 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

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