Submitted by the experts from Germany

Informal document **GRVA-12-03**
12th GRVA, 24 – 28 January 2022
Agenda item 3

Input on A.I. (based on GRVA-03-11)

Germany welcomes that Artificial Intelligence (AI) in vehicle technology is gaining importance in GRVA. The first summary document provided by the leadership of GRVA (GRVA-11-03) is a good impulse and starting point to promote the exchange in the field of AI in automotive engineering. There are particular technical and legal issues in this new field, which need careful consideration.

Germany herewith provides initial comments to GRVA-11-03 and looks forward to a continued exchange with experts in GRVA. Germany is furthermore willing to join additional discussions (e.g. special workshop or dedicated working group) in order to adequately discuss and accompany this topic in the context of GRVA.

Artificial Intelligence and Vehicle Regulations

I. Context

A. Technological developments

1. Artificial Intelligence (AI) has found some prominent applications in the automotive sector. Some of these applications are related to infotainment and vehicle management (as Human Machine Interface (HMI) enhancement) e.g. infotainment management (incl. destination entry in the navigation systems) including voice assistants, which are software agents that can interpret human speech and respond with a synthesized voice. Some applications are related to the development of the safety critical functions (including active safety features, Advanced Driver Assistance Systems and Automated Driving Systems).

B. Mandate

Documentation: ECE/TRANS/WP.29/GRVA/10, para. 18
Informal document GRVA-08-10

2. The Working Party on Automated/Autonomous and Connected Vehicles (GRVA) received inputs regarding Artificial Intelligence in wheeled Vehicles, falling in the scope of the World Forum for Harmonization of Vehicle Regulations.

3. GRVA consulted the Administrative Committee on the Coordination of Work (WP.29/AC.2) on this matter in Fall 2020 and the Committee confirmed that WP.29 considers Artificial Intelligence as an important topic. The Committee raised the question of the need to develop a specific Resolution. It decided, for the time being, to request that GRVA continues addressing this item, also with the aim to develop definitions first and then, corresponding requirements in the scope of WP.29 activities, if necessary (see GRVA-08-10).

4. Following discussions at GRVA level, GRVA requested the Secretary, together with the Officers of the group, to draft a document summarizing the situation and proposing a way forward (ECE/TRANS/WP.29/GRVA/10, para. 18).

II. Positions expressed at the eighth, ninth and tenth sessions of GRVA in 2020 and 2021

A. Definitions

Documentation: Informal documents WP.29-175-21, GRVA-09-23

5. The expert from the Russian Federation suggested that GRVA reviews the definitions provided in ISO/IEC 22989:2020.

6. Several experts noted the importance to not focus on all types of Artificial Intelligence as it would be too broad and not always relevant for the World Forum.

7. The AI related technologies currently used in series vehicles suggest that the Artificial Intelligence at stake in this discussion is limited to the kind of machine learning algorithm used to produce a software with a stochastic approach on the basis of data, leading to complex life cycles and qualitatively new challenges ~~limited to the kind of algorithm used to produce a software on the basis of data and with a stochastic approach.~~ (~~i~~It includes as an example neural networks[1] trained with data.). The nature of the technology involved with its embedding in a noisy sensory-motor-loop implies that ~~technology involved implies that~~ outcomes generated by these processes have a probabilistic nature; building software in that way is not primarily based the usual deterministic logic, it is built using probabilistic reasoning algorithms, based on data~~observation~~ and statistics.

8. There are two types of use that can be distinguished:

(a) The use of this technology can be for the purpose of developing a software in the system development phase; or

(b) The use of the technology is made in service, in order to train i.e. improve the quality of the system used in operation.

Note: it seems that the key points that define the AI based systems of relevance for WP.29 are:

(a) That quality and quantity of data have an influence for training and testing.~~The probabilistic nature of the technology; and~~

(b) The black box approach of AI.~~potential training features belonging to the technology.~~

(c) That output variables can be derived from predictions.

(d) That allows to solve problems that were previously not solvable by classical IT methods and allows to improve the system's performance.

B. Specific input in the Type Approval context

9. The expert from Norway mentioned the readiness of the existing regulatory framework, which includes UN Regulation No. 156 (Software Updates and Software Update Management Systems), providing a useful basis to address the thematic of the software evolution supported by AI and also framing the use of "self-training features".

[1] See WP.29-175-21 (https://unece.org/DAM/trans/doc/2018/wp29/WP29-175-21e.pdf), page 2.

Kommentiert [D1]: Definitions should be consistent when synonyms are used (yellow markings).

A limitation to probabilistic (stochastic) algorithms, models, methods or systems is not realistic in automotive engineering.

Kommentiert [D2]: Explanation: AI generated software, including DNNs, is often deterministic per se but display a stochastic behaviour due to a high input sensitivity in combination with sensory noise and a high dimensional input space (as e.g. CCD camera data).

C. Safety considerations on the use of online learning~~self-learning features~~

10. The expert from the International Telecommunication Union advised that the use of AI agents should be limited to the development phase. He described the suitable use as follow:

(a) An AI agent may be trained to produce a software.

(b) Once satisfactory results are reached, the software should be frozen.

(c) The frozen software should then be validated.

(d) Once properly validated, such software can be placed in vehicle.

11. The expert from the Russian Federation also stated that self-learning functions in operation should not be allowed.

12. There is a continuous transition between a static AI system and an AI system which is retrained online on the edge, i.e. in the vehicle. In between mid-frequency offline retraining combined with a thorough validation and OTA updates offers a compromise that allows adaptations to model drift and model staleness processes while guaranteeing a certain level of safety and security.

Kommentiert [D3]: The closed methodology has an advantage concerning security, but the disadvantage of not always being able to accurately represent the real world.

Kommentiert [D4]: Clarification concerning the terms self-learning / online learning and frequent retraining combined with OTA updates.

D. Technology neutrality, as a best practice for regulations

12. The expert from the Russian Federation recalled that the best practices in terms of regulation is to develop technology neutral provisions.

E. Inherent risks posed by the technology with regards to safety

13. GRVA already faced the situation where the probabilistic nature of systems regulated had to be addressed in order to define suitable performance requirements.

14. Notably, Advanced Emergency Braking Systems such as those regulated by UN Regulation No. 152 and which may be developed using for example Machine Learning or Deep Learning based algorithms or any other technology having a probabilistic nature for the object detection and response, required GRVA to adopt specific provisions as reflected in para. 6.10. (Robustness).

F. Possible ways to address current challenges

15. The expert from CLEPA, having in mind well known risks and challenges associated with data and AI (See III/B below), mentioned that such risks could be covered by the audits performed e.g. in the context of the annexes on complex electronic systems (CEL) (See UN Regulation No. 13, 13-H, 79, 152 and 157).

Kommentiert [D5]: These systems are not auditable to a sufficient degree using existing auditing strategies and tools. New strategies and tools, which are the subject of current research and development activities, are needed.

Many problems and vulnerabilities that are new to data driven AI systems with complex life cycles have not been solved yet. Auditing and problem mitigation strategies have to consider the whole life cycle of the AI system.

Kommentiert [D6]: The methodology used in these UN Regulations is not sufficient for auditing an AI system.
Suggestion: Independence and randomization of test cases and test data, as well as the qualification of the auditor are of paramount importance.

III. Other views expressed

A. Defeating the purpose of regulations

16. The representative of the United Kingdom of Great Britain and Northern Ireland mentioned at WP.29 the potential risk that a self-learning algorithm could potentially

behave in a way that would be comparable to cycle beating or be a defeat device, as it would learn to respond e.g. to a regulated test cycle.

Kommentiert [D7]: The behavior described here depends on the training and the testing data of the AI system.

B. Specific risks associated with the use of data to develop a probabilistic algorithm

17. Several prominent cases were reported that demonstrated the potential risks associated to the use of such technology[2, 3, 4, 5].

18. The quality and quantity of the data has an influence on the bias affecting software built using methods of AI.

19. Having in mind some of the challenges mentioned above, the expert from CLEPA mentioned at GRVA that such risks could be covered~~2~~ by the audits performed e.g. in the CEL annexes of UN Regulation No. 13, 13-H, 79, 152 and 157.

Kommentiert [Germany8]: See remark to above item F. 15.

IV. Possible way forward in terms of committees' activities

A. The role of GRVA

20. The impact of AI on vehicle regulations go beyond the current activities of GRVA dealing with Automated Driving Systems, Advanced Driver Assistance System, Active Safety features and connected vehicles.

21. An example of AI system in vehicles is the Human Machine Interface performing speech recognition and interacting with the driver with regards to the command of head units (incl. navigation systems, air conditioning etc.)

22. GRVA has been tasked to deal in a first step with AI as it is the group that focus most on digital aspects of technology in vehicles.

23. After the Chair of GRVA reported to WP.29 (ECE/TRANS/WP.29/1159, para. 66) the representative of OICA explained that some contracting parties were initiating regulatory activities having relevance for automated driverless shuttles and that others were launching regulatory activities on Artificial Intelligence.

24. He proposed that WP.29 reflects on the harmonization of these matters. GRVA may wish to contribute to this reflection.

B. The role of other groups

25. GRVA has a role to play to address AI in vehicles when it is about safety, ADAS, ADS and connectivity. But other subsidiary bodies of WP.29 might have to deal in future with the specificities of the technologies and the new aspects impacting its work, which may include, among others, data considerations.

[2] https://eu.usatoday.com/story/tech/2015/07/01/google-apologizes-after-photos-identify-black-people-as-gorillas/29567465/

[3] https://nypost.com/2017/12/21/chinese-users-claim-iphone-x-face-recognition-cant-tell-them-apart/

[4] https://docs.house.gov/meetings/GO/GO00/20200115/110380/HHRG-116-GO00-20200115-SD004.pdf

[5] https://spectrum.ieee.org/tech-talk/artificial-intelligence/machine-learning/in-2016-microsofts-racist-chatbot-revealed-the-dangers-of-online-conversation

26. Before the inception of GRVA, WP.29 tasked the IWG on ITS to deal with the initial considerations related to automated and connected vehicles. It should be discussed if a similar approach could be chosen was WP.29 with regards to relevant aspects related to AI and Vehicles.

V. Possible way forward in terms of substantial work; Proposal for a Guidance addressing current known issues

27. As harmonization of technical regulations in that field is premature, WP.29/GRVA may wish to develop guidance on that matter, in the similar way it has been done on cyber security (See ECE/TRANS/WP.29/2017/46 adopted before UN Regulation No. 155)

[Draft Guidance document on the use of AI in vehicles]

The [Member States], [contracting parties to the 1958 and the 1998 Agreements], participating in the Working Party on Automated/Autonomous and Connected Vehicles,

Having recognized the significant penetration of some so-called Artificial Intelligence based systems in wheeled vehicles covered in the scope of the agreements administered by the World Forum for Harmonization of Vehicle Regulations,

Having discussed the technical fundamental aspects of some of the systems in automotive products, which are belonging to what the general public calls Artificial Intelligence, i.e. systems developing software with a probabilistic nature using data, it becomes obvious that the use of these systems entails qualitatively new challenges,

Having assessed the potential incompatibility of *self-learning systems when the vehicle is in operation* with the existing regulatory regimes and with the safety expectations,

Having noted the potential risks that self-learning functions respond in an inadequate manner to test cycles and testing procedures that would lead to defeat device and cycle beating,

Having agreed that it would be premature to harmonize technical regulations on Artificial Intelligence as it would probably impact innovation,

Having acknowledged that the technology was still in developments,

Having discussed that recording and keeping relevant data would support investigations, if need arise, and would help to learn and to inform future regulatory developments,

Have agreed on the following guidance to the stakeholders employing so called AI agent(s) to develop their products:

Online learning~~**Self-learning features**~~

1. This guidance document recalls the existing miles stones of regulatory regimes including, among others, type approval, self-certification, market surveillance, Conformity of Production, In Use Testing, Periodic Technical Inspections and highlights the importance that software versions are kept consistently in all vehicles of a vehicle type.

2. Therefore, the stakeholders should not implement online learning ~~self-learning features~~

(a) which learn during the operation of the vehicle; and/or~~which learn during the operation of the vehicle; and/or~~

Kommentiert [D9]: By setting quality requirements for new technology innovation and trust in new products can be promoted. In addition, requirements can give manufacturers legal security for product development.

Kommentiert [D10]: A large amount of data is addressed. For the implementation, certain areas must be defined (tools for data storage, duration of data storage, location of data storage, scope of data storage).

(b) leading to a situation where two vehicles of the same type differ in their retrained AI models~~self-learning features~~.

3. It is recommended, after having trained an AI-agent to build a software, to freeze that software, and to validate and assess that software with regards to safety and other relevant requirements. Following that process, the validated software may be employed and rolled out in vehicles of a vehicle type. For the avoidance of doubt the frozen software can be updated.

Training data

4. The stakeholders using data to train an AI-agent for the purpose of developing software, including software with a probabilistic nature, should keep record of the data used in training and testing.

5. The stakeholders should ensure that they can provide information in case of inquiry and procedural rights.

6. The stakeholders should implement specific validation methods (regarding safety and other relevant requirements) for systems using stochastic algorithms.

7. The stakeholders should consider and implement possibilities for updates and retraining, to be able to address concerns, non-conformity rectification, recall orders etc.

8. The stakeholders should verify their data in terms of ethics, data protection and privacy, and other general requirements applicable to data in the markets where their product could potentially be used. In the context of the 1958 Agreement, such evaluations can be performed in the framework of the complex electronic systems (CEL) Annexes (e.g. Annex 6 to UN Regulation No. 79).

Black box effect

9. The stakeholders should implement ~~sufficient human~~ state of the art supervision in their processes so that the stakeholders understand the functioning of the software before it is rolled out. ~~Feasibility and traceability depend on the AI model and its complexity (e.g. for DNN this is difficult to achieve). They should be able to explain all functional and safety relevant aspects of it.~~]

Kommentiert [D11]: Situation depending on the used software and update policy.

Kommentiert [Germany12]: We would like to suggest to include a focus on safety relevant assessment and validation methodology in future discussions on the matter.

Kommentiert [D13]: Improves traceability, but a large amount of data is addressed (see above).

Kommentiert [D14]: Feasibility and traceability depend on the AI model and its complexity (e.g. for DNN this is difficult to achieve). Explainable AI methods might mitigate it to some extent but, according to state of the art research, not completely. Human supervision on its own will not be helpful to achieve this goal.