



# Conseil économique et social

Distr. générale  
10 juillet 2023  
Français  
Original : anglais

---

## Commission économique pour l'Europe

Comité des transports intérieurs

### Forum mondial de l'harmonisation des Règlements concernant les véhicules

Groupe de travail des véhicules automatisés/autonomes et connectés

#### Dix-septième session

Genève, 25-29 septembre 2023

Point 3 de l'ordre du jour provisoire

**Véhicules et intelligence artificielle**

## **Proposition de projet de résolution visant à fournir des orientations sur l'intelligence artificielle dans le contexte des véhicules routiers**

### **Communication de l'expert de l'Organisation internationale des constructeurs d'automobiles (OICA)\***

Le texte ci-après, établi par l'expert de l'Organisation internationale des constructeurs d'automobiles (OICA), est fondé sur le document informel GRVA-16-48. Il vise à fournir des orientations sur l'intelligence artificielle dans le contexte des véhicules routiers.

---

\* Conformément au programme de travail du Comité des transports intérieurs pour 2023 tel qu'il figure dans le projet de budget-programme pour 2023 (A/77/6 (Sect. 20), par. 20.6), le Forum mondial a pour mission d'élaborer, d'harmoniser et de mettre à jour les Règlements ONU en vue d'améliorer les caractéristiques fonctionnelles des véhicules. Le présent document est soumis en vertu de ce mandat.



## I. Mandat

1. Le Groupe de travail des véhicules automatisés/autonomes et connectés (GRVA) a reçu des contributions relatives à l'intelligence artificielle appliquée aux véhicules à roues qui relèvent du champ de compétences du Forum mondial de l'harmonisation des règlements concernant les véhicules (WP.29).
2. Le GRVA a consulté le Comité de gestion pour la coordination des travaux (AC.2) sur cette question à l'automne 2020. Celui-ci s'est interrogé sur la nécessité d'élaborer une résolution spécifique. Il a demandé au GRVA de continuer d'étudier ce point pour le moment afin d'établir dans un premier temps des définitions puis, si nécessaire, des prescriptions correspondantes dans le cadre des activités du WP.29 (voir GRVA-08-10).

## II. Proposition

### Résolution sur l'intelligence artificielle dans le contexte des véhicules routiers

#### Préambule

Les [États membres], [Parties contractantes aux Accords de 1958 et de 1998] participant aux travaux du Groupe de travail des véhicules automatisés/autonomes et connectés,

*Ayant constaté* la place toujours plus importante de l'intelligence artificielle dans les véhicules à roues relevant du champ d'application des accords administrés par le Forum mondial de l'harmonisation des Règlements concernant les véhicules (WP.29) [et passé en revue les principaux cas d'utilisation mentionnés à l'Annexe 2],

*Ayant relevé* que le secteur pourrait actuellement utiliser les algorithmes de l'apprentissage automatique pour créer des versions gelées des logiciels et/ou pour les processus d'essais automatisés,

*Ayant examiné* les aspects techniques fondamentaux de certains systèmes d'apprentissage automatique intégrés dans les produits automobiles, que le grand public connaît sous le nom d'intelligence artificielle, et réfléchi aux définitions correspondantes [jointes en annexe au présent document],

*Rappelant* l'adoption des Recommandations relatives à des prescriptions uniformes concernant la cybersécurité et les mises à jour logicielles (ECE/TRANS/WP.29/2022/60) et du Règlement ONU n° 156 (Mises à jour logicielles et système de gestion des mises à jour logicielles),

*Ayant réfléchi* à l'importance que revêt le caractère adéquat des cycles de vie de l'intelligence artificielle pour la compatibilité avec les régimes réglementaires existants et les attentes en matière de sécurité,

*Conscients* du fait que la technologie est encore en évolution,

*Ont arrêté ensemble* les recommandations suivantes relatives à l'utilisation d'algorithmes basés sur l'intelligence artificielle dans leurs produits automobiles :

#### Versions des logiciels

1. Le présent document d'orientation s'applique aux prescriptions réglementaires relatives à l'homologation de type, à l'autocertification, à la surveillance du marché, à la conformité de la production et aux contrôles techniques périodiques. Le secteur ne devrait pas publier de mises à jour logicielles qui modifient une fonction déjà certifiée ou qui influent sur elle, sans recommencer la procédure de certification correspondante.

2. Dans le contexte de l'Accord de 1958 et de l'homologation de type, cela signifie que le secteur ne devrait pas publier de mises à jour logicielles qui modifient une fonction homologuée ou qui influent sur elle, sans consulter de nouveau l'autorité d'homologation. Les mises à jour logicielles qui n'ont pas d'incidence sur la performance des fonctions homologuées peuvent être déployées sans qu'il soit nécessaire de reconsulter l'autorité d'homologation.

#### Cycle de vie de l'intelligence artificielle

3. Il est recommandé de ne geler une version d'un logiciel qu'après avoir entraîné le système d'intelligence artificielle qui sera intégré à ce logiciel. Cette version doit être validée et évaluée au regard des prescriptions en matière de sécurité et des autres prescriptions pertinentes. À l'issue de ce processus, le logiciel validé peut être déployé dans des véhicules d'un type donné.

#### Données d'entraînement

3. Il est recommandé de vérifier les données utilisées du point de vue de la protection des données et de l'intimité de la vie privée ainsi que d'autres obligations légales. La présente résolution s'entend sans préjudice des dispositions législatives et réglementaires propres au marché qui régissent la manière dont les données personnelles sont collectées et utilisées. Lorsque de telles réglementations existent, elles contribuent à la sécurité globale du système d'intelligence artificielle en fixant des normes de sécurité en matière de gestion des données personnelles.

## Annexe 1

### Définitions simplifiées applicables dans le contexte des règlements concernant les véhicules

Les termes ci-dessous sont en grande partie tirés des définitions qui sont à l'étude au sein de l'Organisation internationale de normalisation (ISO) (voir ISO/IEC 22989).

- **Agent** : tout élément qui perçoit son environnement, agit de manière autonome pour atteindre des objectifs, et peut améliorer son efficacité grâce à l'apprentissage ou utiliser des connaissances.
- **Cycle de vie de l'intelligence artificielle** : phase de conception et de développement du système basé sur l'intelligence artificielle, notamment, la collecte, la sélection et le traitement des données, le choix du modèle et le processus d'entraînement, la phase de validation, la phase de déploiement et la phase de surveillance. Ce cycle prend fin lorsque le système basé sur l'intelligence artificielle n'est plus opérationnel.
- **Intelligence artificielle** : ensemble de méthodes ou d'entités automatisées qui, ensemble, construisent, optimisent et appliquent un modèle permettant au système, pour un ensemble donné de tâches prédéfinies, de calculer des prédictions, des recommandations ou des décisions.
- **Biais** : différence systématique dans le traitement (y compris la catégorisation/l'observation) de certains objets (par exemple des personnes physiques ou des groupes) par rapport à d'autres.
- **Boîte noire** : système/logiciel dont l'architecture détaillée et le traitement sont inconnus.
- **Tests en boîte noire/grise/blanche** : tests de systèmes ou logiciels dont l'architecture et le traitement sont inconnus/partiellement connus/connus.
- **Systèmes d'intelligence artificielle connexionnistes** : systèmes généralement constitués de nombreux nœuds, appelés neurones, qui sont interconnectés selon des structures particulières, en fonction du modèle d'intelligence artificielle utilisé. Les réseaux neuronaux et les machines à vecteurs de support sont des exemples de systèmes d'intelligence artificielle connexionnistes. Pour de nombreux cas d'utilisation, les systèmes d'intelligence artificielle de type connexionniste s'avèrent plus puissants que ceux de type symbolique, par exemple dans le domaine de la vision par ordinateur. Dans la majorité des cas, les paramètres des systèmes d'intelligence artificielle connexionnistes ne peuvent pas être définis directement par le développeur. Ce sont des algorithmes d'apprentissage automatique couplés à des données qui sont utilisés à la place pour entraîner ces systèmes. La qualité du système d'intelligence artificielle connexionniste qui en résulte dépend essentiellement de la qualité et de la quantité des données d'entraînement. Contrairement aux systèmes d'intelligence artificielle symboliques, les systèmes connexionnistes sont, dans la plupart des cas, difficilement interprétables et non vérifiables formellement.
- **Annotation des données** : processus qui consiste à associer un ensemble d'informations descriptives à des données sans modifier ces dernières.
- **Échantillonnage des données** : processus statistique qui consiste à sélectionner un sous-ensemble de données devant faire ressortir des formes et des tendances semblables à celles mises en évidence grâce à l'analyse de l'ensemble complet de données.
- **Ensemble de données** : jeu de données qui ont le même format et dont le contenu est pertinent pour l'objectif.
- **Apprentissage profond** : processus par lequel les réseaux neuronaux utilisent plusieurs couches de traitement devant permettre d'extraire progressivement à partir des données des caractéristiques de niveau supérieur.

- **Explicabilité** : propriété d'un système basé sur l'intelligence artificielle qui permet d'exprimer les facteurs importants influençant le résultat du système d'une manière compréhensible pour l'être humain.
- **Équité/matrice d'équité** : une des façons de décrire les biais.
- **Boîte grise** : système/logiciel dont les détails de l'architecture et du traitement sont partiellement connus.
- **Contrôle humain** : propriété du système basée sur l'intelligence artificielle qui garantit que les contraintes opérationnelles intégrées ne peuvent pas être ignorées par le système lui-même et qu'elles tiennent compte de l'opérateur humain, et que les personnes physiques auxquelles la supervision humaine est confiée exercent le contrôle en dernier lieu.
- **Apprentissage automatique** : ensemble de méthodes de calcul fondées sur des données qui permettent de créer la capacité d'apprendre sans suivre d'instructions explicites, de sorte que le comportement du modèle fait ressortir des formes dans les données ou l'expérience.
- **Modèle d'apprentissage automatique** : construction informatique qui génère une inférence, ou une prédiction, basée sur des données d'entrée.
- **Dérive du modèle** : terme utilisé dans le domaine de l'apprentissage automatique. Il désigne le phénomène selon lequel la précision des prédictions des modèles d'apprentissage automatique peut se dégrader au fil du temps. Les raisons en sont, par exemple, que les hypothèses ou les dépendances des variables qui étaient encore valables lorsque les modèles ont été créés et entraînés ont évolué au fil du temps. Des mesures telles que le réentraînement ou l'ajustement des modèles peuvent permettre d'éliminer les dérives.
- **Obsolescence du modèle** : il y a obsolescence du modèle entraîné lorsque celui-ci ne contient pas de données fraîches et/ou ne répond pas aux prescriptions actuelles. Les modèles obsolètes peuvent compromettre la qualité de la prédiction dans les logiciels intelligents.
- **Apprentissage en ligne** : entraînement progressif d'une nouvelle version d'un système basé sur l'intelligence artificielle pendant le fonctionnement de ce système à bord d'un véhicule de série en vue d'atteindre des objectifs précis.
- **Prévisibilité** : propriété d'un système basé sur l'intelligence artificielle qui permet aux parties prenantes de formuler des hypothèses fiables sur les résultats.
- **Apprentissage par renforcement** : discipline de l'apprentissage automatique qui permet à un agent d'apprendre à prendre des décisions à partir de formes dans les données ou les expériences, en maximisant une récompense quantitative au fil du temps.
- **La fiabilité** : propriété qui permet d'obtenir un comportement et des résultats cohérents.
- **Résilience** : capacité d'un système à retrouver rapidement son état de fonctionnement après un incident.
- **Robustesse** : capacité d'un système à maintenir son niveau de performance dans toutes sortes de circonstances.
- **Sécurité dès la conception** : propriété du système rendue possible grâce à des actions en amont relatives au développement et au cycle de vie, qui vise à garantir que les risques sont ramenés à un niveau acceptable grâce aux mesures prises concernant ce système.
- **Apprentissage semi-supervisé** : apprentissage combinant les apprentissages supervisé et non supervisé. Il requiert une petite quantité de données étiquetées et une grande quantité de données non étiquetées, ce qui permet de bénéficier des avantages à la fois de l'apprentissage non supervisé et de l'apprentissage supervisé tout en évitant les difficultés liées à la recherche d'une grande quantité de données étiquetées.

- **Logiciel** : en général, fruit d'un processus appelé programmation traditionnelle. Le programmeur code manuellement les règles en utilisant un langage de programmation.
- **Apprentissage supervisé** : type d'apprentissage automatique qui utilise des données étiquetées pendant l'entraînement.
- **Intelligence artificielle symbolique** : les systèmes d'intelligence artificielle symbolique encodent explicitement les connaissances à l'aide de représentations symboliques. Les arbres de décisions sont un exemple de ce type de système. Il est généralement possible et beaucoup plus facile d'interpréter et de vérifier formellement un système de type symbolique qu'un système de type connexionniste.
- **Entraînement** : processus qui vise à ajuster les paramètres d'un modèle d'apprentissage automatique.
- **Données d'entraînement** : sous-ensemble des échantillons de données d'entrée utilisés pour entraîner un modèle d'apprentissage automatique.
- **Transparence d'une organisation** : propriété d'une organisation selon laquelle les activités et les décisions appropriées sont solidement étayées et communiquées aux parties prenantes concernées de manière exhaustive, accessible et compréhensible.
- **Transparence d'un système** : propriété d'un système qui permet de communiquer des informations aux parties prenantes.
- **Crédibilité** : capacité à répondre aux attentes des parties prenantes de manière vérifiable.
- **Apprentissage non supervisé** : type d'apprentissage automatique qui utilise des données non étiquetées pendant l'entraînement.
- **Validation** : permet de s'assurer de la facilité d'utilisation du logiciel et de sa capacité à répondre aux besoins du client.
- **Données de validation** : données utilisées pour évaluer l'efficacité d'un modèle final d'apprentissage automatique.
- **Vérification** : opération effectuée dans le but de s'assurer que le logiciel est de bonne qualité, bien conçu, robuste et exempt d'erreurs, sans répercussions sur sa facilité d'utilisation.
- **Boîte blanche** : système/logiciel dont l'architecture détaillée et le traitement sont connus.

## Annexe 2

## Examen des cas d'utilisation dans les véhicules communiqués par le secteur en 2022

Application de l'intelligence artificielle		Fonctions non liées à la sécurité Par exemple infodivertissement Hors du champ d'application de l'homologation de type	Fonctions liées à la sécurité				
			Fonction de conduite			Fonctions non liées à la conduite	
			Perception	Planification	Commande		
Logiciel classique	<b>Intelligence artificielle</b>	L'intelligence artificielle est un ensemble de méthodes ou d'entités automatisées qui, ensemble, construisent, optimisent et appliquent un modèle permettant au système, pour un ensemble donné de tâches prédéfinies, de calculer des prédictions, des recommandations ou des décisions.	Traitement du langage naturel	Hors du champ d'application [Pas d'intelligence artificielle]  Détection des autres usagers de la route pour systèmes AEBS, régulateurs ACC Détection des infrastructures routières pour systèmes LDWS et LKAS	Hors du champ d'application  Activation des systèmes FCW et AEBS en fonction de la position du véhicule soumis à l'essai et des autres usagers de la route	Sans objet	Hors du champ d'application  Détection du visage du conducteur pour l'identification (dans des conditions garantissant le respect de la vie privée)
	<b>Apprentissage supervisé</b>	L'apprentissage supervisé est un type d'apprentissage automatique qui utilise des données étiquetées pendant l'entraînement.	Commande gestuelle Reconnaissance vocale	Détection des autres usagers de la route pour AEBS, régulateurs ACC Détection des infrastructures routières passives pour systèmes LDWS et LKAS	Prédiction de trajectoire à partir de données étiquetées (par exemple, cartes HD)	Sans objet	Détection du regard/de l'état du conducteur pour le système de surveillance du conducteur. Détection des défaillances, Maintenance prévue

Intelligence artificielle	<b>Apprentissage non supervisé</b>	L'apprentissage non supervisé est un type d'apprentissage automatique qui utilise des données non étiquetées pendant l'entraînement.		Rationalisation du processus d'étiquetage des données pour les systèmes moins critiques pour la sécurité, tels que les ISA. Extraction de scénarios à partir de données du monde réel pour appuyer la validation. Génération de données de synthèse pour l'apprentissage supervisé/distorsion des données du monde réel	Prédiction de trajectoire à l'aide de filtres Kalman, architectures KalmanNet, gaussiennes ou autres	Sans objet	[?]
	<b>Apprentissage semi-supervisé</b>	L'apprentissage semi-supervisé est une technique qui permet d'« apprendre » à partir d'un mélange de données étiquetées et de données non étiquetées et non structurées. Il s'appuie sur un petit ensemble d'exemples connus et utilise ensuite ces informations pour guider l'apprentissage non supervisé.		Rationalisation du processus d'étiquetage des données pour les systèmes moins critiques pour la sécurité, tels que les ISA.	La technique dite « shadow mode » est utilisée dans le cadre du développement d'algorithmes de contrôle de l'entraînement	Sans objet	[?]
	<b>Apprentissage par renforcement</b>	L'apprentissage par renforcement est un type d'apprentissage automatique qui utilise une fonction de récompense pour optimiser un modèle d'apprentissage automatique par interaction séquentielle avec un environnement.		Certains fabricants commencent à utiliser l'apprentissage par renforcement pour la perception ; celui-ci pourrait être utilisé dans le cadre de la perception coopérative à l'avenir.	L'apprentissage par renforcement peut être utilisé pour les systèmes de centrage de voie ou ACC compte tenu de la réduction du coût et des données nécessaires à l'entraînement du système.	Sans objet	Maintenance prédictive