|  |  |  |
| --- | --- | --- |
|  | Организация Объединенных Наций | ECE/TRANS/WP.29/GRVA/2023/17 |
| _unlogo | **Экономический и Социальный Совет** | Distr.: General10 July 2023RussianOriginal: English |

**Европейская экономическая комиссия**

Комитет по внутреннему транспорту

**Всемирный форум для согласования правил
в области транспортных средств**

**Рабочая группа по автоматизированным/автономным и подключенным транспортным средствам**

**Семнадцатая сессия**

Женева, 25–29 сентября 2023 года

Пункт 3 предварительной повестки дня

**Искусственный интеллект в транспортных средствах**

 **Предложение по проекту резолюции с руководством по искусственному интеллекту в контексте дорожных транспортных средств**

 **Представлено экспертом от Международной организации предприятий автомобильной промышленности**[[1]](#footnote-1)\*

 Воспроизведенный ниже текст был подготовлен экспертом от Международной организации предприятий автомобильной промышленности (МОПАП), и в его основу положен неофициальный документ GRVA-16-48. Он призван обеспечить руководство по искусственному интеллекту (ИИ) в контексте дорожных транспортных средств.

 I. Мандат

1. Рабочей группой по автоматизированным/автономным и подключенным транспортным средствам (GRVA) были получены материалы, касающиеся использования искусственного интеллекта в колесных транспортных средствах, относящихся к компетенции Всемирного форума для согласования правил в области транспортных средств (WP.29).

2. Осенью 2020 года GRVA провела консультации по этому вопросу с Административным комитетом по координации работы (WP.29/AC.2). Комитет поднял вопрос о необходимости разработки специальной резолюции. На тот момент им было решено обратиться к GRVA с просьбой продолжить рассмотрение данного вопроса, в том числе с целью разработки сперва определений, а затем — при необходимости — и соответствующих требований в контексте деятельности WP.29 (см. документ GRVA-08-10).

 II. Предложение

«**Резолюция по искусственному интеллекту в контексте дорожных транспортных средств**

Преамбула

[Государства-члены], [Договаривающиеся стороны Соглашений 1958 и 1998 годов], участвующие в Рабочей группе по автоматизированным/автономным и подключенным транспортным средствам,

 *признавая* значительную степень проникновения некоторых видов искусственного интеллекта (ИИ) в колесные транспортные средства, подпадающие под действие соглашений, относящихся к ведению Всемирного форума для согласования правил в области транспортных средств (WP.29) [, и рассмотрев основные варианты использования, приведенные в приложении 2],

 *принимая во внимание*, что в настоящее время отрасль может использовать алгоритмы машинного обучения для создания подконтрольных версий программного обеспечения и/или для поддержки автоматизированных процессов тестирования,

 *обсудив* основополагающие технические аспекты некоторых систем машинного обучения, используемых в автомобильной продукции, которые широкая публика называет искусственным интеллектом, и рассмотрев соответствующие определения [, прилагаемые к настоящему документу],

 *напоминая* о принятии рекомендаций по единообразным предписаниям, касающимся кибербезопасности и обновления программного обеспечения (ECE/TRANS/WP.29/2022/60), и Правил № 156 ООН (обновление программного обеспечения и система управления обновлениями программного обеспечения),

 *проведя оценку* важности надлежащих жизненных циклов ИИ для обеспечения совместимости с существующими нормативными режимами и ожиданиями в плане безопасности,

 *признавая* тот факт, что данная технология все еще находится в стадии разработки,

 *согласовали* нижеследующие рекомендации по использованию алгоритмов на базе ИИ в своей автомобильной продукции:

Версии программного обеспечения

1. Настоящее руководство распространяется на нормативные требования, предъявляемые к официальному утверждению типа, самосертификации, наблюдению за рынком, контролю за соответствием производства и периодическим техническим осмотрам. Отрасль не должна выпускать обновления программного обеспечения, которые изменяют или влияют на уже сертифицированную функцию, без возобновления соответствующей процедуры сертификации.

2. В контексте Соглашения 1958 года и официального утверждения типа это означает, что отрасль не должна выпускать обновления программного обеспечения, которые изменяют или влияют на функцию, официально утвержденную для данного типа, без повторного согласования с органом по официальному утверждению. Обновления программного обеспечения, не влияющие на характеристики, официально утвержденные для данного типа, могут внедряться без повторного согласования с органом по официальному утверждению.

Жизненный цикл искусственного интеллекта

3. Рекомендуется — по завершении обучения системы на базе ИИ — ставить под контроль версию, заложенную в программное обеспечение. Она подлежит валидации и оценке с точки зрения безопасности и других соответствующих требований. По завершении этого процесса валидированное программное обеспечение может развертываться на транспортных средствах того или иного типа.

Обучающие данные

3. Рекомендуется проверять используемые данные с точки зрения защиты и конфиденциальности информации, а также на соответствие другим законодательным требованиям. Настоящая резолюция не наносит ущерба действующим на конкретном рынке законодательным и нормативным актам, касающимся порядка сбора и использования персональных данных. Там, где такие нормы существуют, они способствуют повышению общей безопасности системы ИИ, устанавливая стандарты безопасности управления персональными данными.

[Приложение 1

 Упрощенные определения в контексте правил в области транспортных средств

 Приведенные ниже термины в значительной мере позаимствованы из определений, находящихся на рассмотрении Международной организации по стандартизации (см. ISO/IEC 22989).

* **Агент** — это сущность, которая воспринимает свое окружение, автономно предпринимает действия для достижения целей и может улучшать свою производительность благодаря обучению либо использовать полученные знания.
* **Жизненный цикл ИИ** состоит из этапа проектирования и разработки системы на базе ИИ, включая, в частности, сбор, отбор и обработку данных, а также выбор модели и процесса обучения; этапа валидации; этапа внедрения; и этапа мониторинга. Жизненный цикл завершается, когда система на базе ИИ перестает функционировать.
* **Искусственный интеллект (ИИ)** — это совокупность методов или автоматизированных сущностей, которые сообща строят, оптимизируют и применяют модель таким образом, чтобы система могла — исходя из заданного набора предварительно определенных задач — вычислять прогнозы, рекомендации или решения.
* **Предвзятость** — это систематическое различие в отношении (включая категоризацию/наблюдение) к определенным объектам (например, людям или группам) по сравнению с другими.
* **«Черный ящик»** — это система/программа, в случае которой детальная архитектура и подробности обработки неизвестны.
* **Тестирование методом «черного»/«серого»/«прозрачного» ящика** — это способы тестирования системы/программы, в случае которой детальная архитектура и подробности обработки неизвестны/частично известны/известны.
* **Коннекционистские ИИ (кИИ-)** системы обычно состоят из множества узлов, называемых нейронами, которые связаны друг с другом по определенным схемам, зависящим от конкретной модели ИИ. Примерами кИИ-систем являются нейронные сети и машины опорных векторов. Во многих прикладных сферах кИИ-системы являются более мощными по сравнению с
сИИ-системами, например в компьютерном зрении. В большинстве случаев параметры кИИ-систем не могут быть заданы непосредственно разработчиком. Вместо этого для обучения таких систем используются алгоритмы машинного обучения в сочетании с соответствующими данными. Качество результирующей кИИ-системы особо зависит от качества и количества обучающих данных. В отличие от сИИ-систем кИИ-системы в большинстве случаев не относятся к числу легко интерпретируемых и не являются формально верифицируемыми.
* **Аннотирование данных** — это процесс присоединения к данным набора описательной информации без какого-либо изменения этих данных.
* **Выборка данных** — это статистический процесс отбора подмножества данных, в котором должны быть представлены закономерности и тенденции, аналогичные тем, которые наблюдаются в более крупном анализируемом наборе данных.
* **Набор данных** — это совокупность данных, имеющих общий формат и содержание, соответствующее конкретной цели.
* **Глубокое обучение** — это процесс, при котором нейронные сети используют большое количество слоев обработки для извлечения из данных все более высокоуровневых характеристик.
* **Объяснимость** означает свойство системы на базе ИИ отражать важные показатели, влияющие на результаты работы системы, в понятном для человека виде.
* **Беспристрастность/матрица беспристрастности** — это способ описания предвзятости.
* **«Серый ящик»** — это система/программа, в случае которой детальная архитектура и подробности обработки известны частично.
* **Человеческий надзор** — это заложенное в систему на базе ИИ свойство, призванное гарантировать, что встроенные операционные ограничения не могут быть отменены самой системой и реагируют на действия человека-оператора и что окончательный контроль осуществляют физические лица, на которых возложена функция человеческого надзора.
* **Машинное обучение** — это совокупность основанных на данных вычислительных методов, обеспечивающих способность системы к обучению без явных инструкций, так чтобы поведение модели отражало закономерности в данных или опыте.
* **Модель машинного обучения** — это компьютерная математическая модель, которая генерирует вывод или предсказание на основе входных данных.
* **Дрейф модели** — это термин из области машинного обучения. Он относится к тому явлению, что с течением времени точность прогнозирования моделей машинного обучения может ухудшаться. Это может быть обусловлено, например, тем, что допущения или зависимости переменных, которые были актуальны на момент создания и обучения моделей, со временем изменились. Устранить дрейф модели позволяют такие меры, как переобучение или настройка моделей.
* **Ветхость модели** определяется как устаревание, если обученная модель не содержит актуальных данных и/или не отвечает современным требованиям. Устаревшие модели могут негативно сказаться на качестве обеспечиваемого интеллектуальным программным обеспечением прогнозирования.
* **Онлайновое обучение** означает постепенное обучение новой версии системы на базе ИИ в процессе эксплуатации на борту серийных транспортных средств для достижения поставленных целей.
* **Предсказуемость** — это свойство системы на базе ИИ, которое позволяет заинтересованным сторонам делать достоверные предположения о результатах работы.
* **Обучение с подкреплением** — это способ машинного обучения, позволяющий агенту обучаться действиям на основе закономерностей в данных или опыте, оптимизируя количественную функцию накопленного за это время вознаграждения.
* **Надежность** — это свойство соответствия ожидаемому поведению и результатам.
* **Отказоустойчивость** — это способность системы к быстрому восстановлению работоспособности после инцидента.
* **Робастность** — это способность системы сохранять свой уровень производительности при самых различных обстоятельствах.
* **Безопасность по умолчанию** — это свойство системы, обеспечиваемое проактивной разработкой и упреждающей деятельностью на протяжении жизненного цикла, позволяющее свести риски к приемлемому уровню за счет системных мер.
* **Обучение с частичным привлечением учителя** — это сочетание способов обучения с учителем и без учителя. При нем используется небольшой объем размеченных данных и большой объем неразмеченных данных, что обеспечивает преимущества обоих способов обучения, позволяя при этом избегать проблем, связанных с поиском большого объема размеченных данных.
* **Программное обеспечение** обычно создается с помощью процесса, называемого традиционным программированием. Программист вручную кодирует правила с использованием языка программирования.
* **Обучение с учителем** — это способ машинного обучения, при котором во время обучения используются размеченные данные.
* **Символьный ИИ (сИИ)** связан с эксплицитным кодированием знаний с использованием символов. Примером такой системы является дерево решений. Интерпретация и формальная верификация сИИ-систем в целом возможны и значительно проще по сравнению с коннекционистскими системами ИИ.
* **Обучение** — это процесс настройки параметров модели машинного обучения.
* **Обучающие данные** — это подмножество образцов входных данных, используемых для настройки модели машинного обучения.
* **Транспарентность организации** — это свойство организации документировать ее деятельность и решения и доносить их до соответствующих заинтересованных сторон в полной, доступной, ясной и понятной форме.
* **Транспарентность системы** — это свойство системы доносить информацию до заинтересованных сторон.
* **Достоверность** — это способность оправдывать ожидания заинтересованных сторон проверяемым способом.
* **Обучение без учителя** — это способ машинного обучения, при котором во время обучения используются неразмеченные данные.
* **Валидация** позволяет убедиться в практичности использования программного обеспечения и его способности отвечать потребностям клиента.
* **Результаты валидации** — это данные, служащие для оценки эффективности конечной модели машинного обучения.
* **Верификация** позволяет убедиться в высоком качестве программного обеспечения, его отлаженности, надежности и отсутствии ошибок, не вдаваясь при этом в практичность его использования.
* **«Прозрачный ящик»** — это система/программа, в случае которой детальная архитектура и подробности обработки известны.

Приложение 2

 Обзор представленных отраслью в 2022 году вариантов использования ИИ в транспортных средствах

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  |  | **Функции,не связанные с безопасностью**(например, информационно-развлекательная)Не подпадают под действие официального утверждения типа | **Функции обеспечения безопасности** |
|  | **Интеллектуальное приложение** | **Функция вождения** | **Функции, не связанныес вождением** |
|  | **Восприятие**  | **Планирование** | **Активация** |  |
| Обычное программноеобеспечение | **Искусственный интеллект (ИИ)** | Искусственный интеллект — это совокупность методов или автоматизированных сущностей, которые сообща строят, оптимизируют и применяют модель таким образом, чтобы система могла — исходя из заданного набора предварительно определенных задач — вычислять прогнозы, рекомендации или решения. | Обработка естественно-языковых текстов | Не подпадает под действие [не относится к ИИ]Обнаружение присутствия других участников дорожного движения для работы систем САЭТ, АСС.Обнаружение объектов дорожной инфраструктуры для работы систем СПВП, СУПП. | Не подпадает под действиеАктивация систем СПФС и САЭТ с учетом положения испытуемого транспортного средства и присутствия других участников дорожного движения. | Неприменимо | Не подпадают под действиеРаспознавание лица водителя для идентификации личности (на условиях обеспечения конфиденциальности). |
|  | **Обучение с учителем (ОсУ)** | Обучение с учителем— этоспособ машинного обучения, при котором во время обученияиспользуются размеченныеданные. | Считываниежестикуляции Распознаваниеголоса | Обнаружение присутствия других участников дорожного движения для работы систем САЭТ, АСС.Обнаружение объектовдорожной инфраструктуры для работы систем СПВП, СУПП. | Прогнозирование траектории движения путем прогнозирования пути перемещения по размеченным данным (например, картам высокого разрешения). | Неприменимо | Распознавание направление взгляда/поворота головы водителя для работысистемы управления данными.Выявление неисправностей, диагностическое обслуживание. |
| Искусственный интеллект | **Обучение без учителя (ОбУ)**  | Обучение без учителя— это способ машинного обучения, при котором во время обучения используются неразмеченные данные. |   | Оптимизация процессаразметки данных для менее критических с точки зрения безопасности систем, таких как ИСА. Извлечение сценариев из реальных данных в обоснование валидации.Генерирование синтетических данных для обучения с учителем/искажения реальных данных. | Прогнозирование траектории движения с использованием фильтров Калмана, архитектур на основе KalmanNet или гауссовского процесса либо других архитектур. | Неприменимо | [?] |
| **Обучение с частичным привлечением учителя (ОчУ)** | Обучение с частичным привлечением учителя— это методика, при которой обучение происходит на «смеси» размеченных данных и данных, которые являются неразмеченными и неструктурированными. За основу берется небольшой набор известных примеров, и затем эта информация используется для обучения без учителя. |   | Оптимизация процесса разметки данных для менее критических с точки зрения безопасности систем, таких как ИСА. | При разработке алгоритмов управления обучением используется режим «тени». | Неприменимо | [?] |
| **Обучение с подкреплением (ОП)** | Обучение с подкреплением — это способ машинного обучения, использующий функцию вознаграждения для оптимизации модели машинного обучения путем последовательного взаимодействия с окружающей средой. |   | Некоторые изготовители начинают использовать ОП для функции восприятия, а в будущем оно вполне может использоваться для целей кооперативного восприятия. | ОП может использоваться в случае систем удержания по центру полосы движения или ACC ввиду снижения стоимости/доступности данных, необходимых для обучения системы. | Неприменимо | Диагностическое обслуживание |

1. \* В соответствии с программой работы Комитета по внутреннему транспорту на 2023 год, изложенной в предлагаемом бюджете по программам на 2023 год (A/77/6 (часть V, разд. 20), п. 20.6), Всемирный форум будет разрабатывать, согласовывать и обновлять правила ООН в целях улучшения характеристик транспортных средств. Настоящий документ представлен в соответствии с этим мандатом. [↑](#footnote-ref-1)