

A futuristic cityscape featuring vertical gardens on building facades and autonomous vehicles on a multi-lane road. The scene is set in a bright, clear sky with modern, cylindrical skyscrapers in the background. The overall aesthetic is clean, green, and technologically advanced.

# ПОДХОДЫ К ПРОБЛЕМЕ ВАГОНЕТКИ

В РАЗРЕЗЕ БЕСПИЛОТНОГО  
НАЗЕМНОГО ТРАНСПОРТА

# Содержание

Введение.....	3
Российский подход.....	4
Международные подходы.....	6

## Введение

Проблема «вагонетки» — это известный философский мысленный эксперимент, который последние сорок лет занимает внимание специалистов по этике, психологов, юристов, философов и инженеров. Традиционный сценарий эксперимента состоит в том, что «сбежавшая» вагонетка движется по пути, на котором находится несколько, как правило, пять человек. Переключив «стрелку», вагонетку можно направить на другие пути, и тогда погибнет только один человек. Все множество других сценариев сводится к одному вопросу: допустимо ли жертвовать одним человеком, ради спасения других жизней?

Эта этическая дилемма выявила разницу между двумя моральным концепциями: осознанное (активное) лишение жизни одного человека ради наивысшей общей пользы - сохранения большего числа жизней или пассивное невмешательство, основанное на принципе «никогда не убивать человека».

С появлением беспилотного транспорта эта дилемма получила новое преломление. На чаше весов оказываются жизни пешеходов и пассажиров, а человеком, который заранее «настраивает» стрелку, становится разработчик алгоритмов искусственного интеллекта.

**” Действительно, такая проблема существует. Можно привести достаточно конкретный пример: система на основе ИИ в экстренной ситуации должна принять решение - сбить пешехода, сохраняя выбранную траекторию движения или из-за резкого маневра подвергнуть опасности пассажира. Но с точки зрения разработчиков, это скорее техническая проблема.**

Илья Попов, к.т.н., доцент факультета безопасности информационных технологий ИТМО

Однако не все разработчики в сфере беспилотного транспорта признают актуальность проблемы «вагонетки».

**” Проблема вагонетки для беспилотного транспорта не совсем актуальна. Беспилотные системы программируют таким образом, чтобы избежать столкновения с внезапно выбежавшим на дорогу человеком в любых условиях: плохой видимости, дождливой погоде и пр. Проблему вагонетки можно создать искусственно. Например, если разработчик сузит зону безопасности или увеличит скорость движения беспилотного транспорта.**

Иван Дейлид, руководитель отдела разработки программного обеспечения Центра беспилотных технологий Университета Иннополис

Существует ли проблема «вагонетки» на практике в сфере беспилотного транспорта? Каким должен быть профессиональный и моральный выбор разработчика? Допустимо ли в принципе закладывать в систему ИИ этический выбор о причинении вреда разным людям (группе людей)? Как должен вести себя беспилотный автомобиль, если столкновения все-таки не избежать?

## Российский подход к проблеме «вагонетки»

По результатам интервью с российскими и зарубежными разработчиками в сфере беспилотного наземного транспорта, а также исследования зарубежных подходов к решению проблемы «вагонетки» Комиссией по этике в сфере искусственного интеллекта с участием экспертов, принявших участие в обсуждении, сформулированы ключевые выводы, принципы и рекомендации. Рекомендации предназначены для широкого круга разработчиков беспилотного наземного транспорта, профессиональных эксплуатантов и обычных пользователей (пассажиров).

### Вывод 1.

На практике проблемы этического выбора – чья жизнь ценнее – при программировании беспилотного транспорта не существует.

### Вывод 2.

Все жизни ценны одинаково.

## Принцип

**Человеческая жизнь – высшая ценность.** Непричинение вреда человеку является главным приоритетом разработчика.

### Этическая «нейтральность»

Никакой этический выбор о причинении (непричинении) вреда не должен закладываться в СИИ.

## Рекомендация

Система искусственного интеллекта (СИИ) должна программироваться так, чтобы избегать опасности причинения вреда человеку, какие бы иные потери ни последовали.

СИИ должна программироваться без предоставления ей права этической оценки рисков причинения вреда и выбора вариантов действий для минимизации последствий.

## Принцип

### **Избежание аварийной ситуации**

Задача программиста не решать аварийную ситуацию. Он должен сделать все, чтобы она не наступила.

## Рекомендация

Рекомендация

СИИ должна программироваться так, чтобы избежать столкновения с человеком в любых условиях (при плохой видимости, дождливой погоде и пр).

Для этого необходимо параметризовать граничные условия среды эксплуатации с учетом различных условий (время суток, погода и пр.): предельно допустимая скорость, коэффициенты сцепления колеса с дорожным покрытием, допустимые ограничения видимости, ограничения дистанции и пр.

### **Отсутствие социальных различий**

Каждая человеческая жизнь ценна одинаково, независимо от возраста, расы, пола, социального статуса и пр.

СИИ не должна программироваться на выбор о причинении (непричинении) вреда человеку по каким-либо параметрам. Задача алгоритма не допустить аварии как таковой.

### **Соблюдение правил дорожного движения**

Все участники дорожного движения, в т.ч. беспилотные автомобили, должны соблюдать правила дорожного движения (ПДД).

СИИ должна программироваться на строгое соблюдение ПДД, включая возможность нарушения ПДД в случаях крайней необходимости для избежания столкновения (снижение/превышение допустимой скорости, нарушение разметки и пр.).

## Зарубежные подходы к проблеме «вагонетки»

По результатам интервью с российскими и зарубежными разработчиками в сфере беспилотного наземного транспорта, а также исследования зарубежных подходов к решению проблемы «вагонетки» Комиссией по этике в сфере искусственного интеллекта с участием экспертов, принявших участие в обсуждении, сформулированы ключевые выводы, принципы и рекомендации. Рекомендации предназначены для широкого круга разработчиков беспилотного наземного транспорта, профессиональных эксплуатантов и обычных пользователей (пассажиров).

- 1. Единственным государственным органом** в мире, который решил заняться решением проблемы «вагонетки», является Федеральное министерство транспорта и цифровой инфраструктуры Федеративной республики Германия. **Комиссия по этике автоматизированного и подключенного вождения**, состоящая из экспертов из широкого спектра дисциплин, включая специалистов в области философии, юриспруденции, социальных наук, оценки воздействия технологий, автомобильной промышленности и разработки программного обеспечения, была созвана федеральным министром транспорта и цифровой инфраструктуры и провела свое учредительное заседание 30 сентября 2016 года. Рабочая группа 1 Комиссии занималась вопросами вреда, которого невозможно избежать (проблемой «вагонетки»)<sup>1</sup>.

Комиссия решила, что решения, в которых стоит выбор между одной человеческой жизнью и другой, зависят от конкретной ситуации, включая «непредсказуемое» поведение затронутых сторон. Таким образом, они не могут быть четко стандартизированы, а устройства не могут быть запрограммированы таким образом, чтобы ситуации были этически бесспорными.

**” В процессе разработки сценариев безопасности нашим коллективом был выделен класс ситуаций, в которых источником угрозы является сам человек. Нарушая правила техники безопасности, например, находясь самовольно в опасных зонах, человек становится основным дестабилизирующим фактором для беспилотного транспортного средства.**

Илья Попов, ИТМО.

---

1. BMVI Commission (28 August 2017). "Ethics Commission's complete report on automated and connected driving". Federal Ministry of Transport and Digital Infrastructure (German: Bundesministerium für Verkehr und digitale Infrastruktur) // URL: <https://bmdv.bund.de/SharedDocs/EN/publications/report-ethics-commission.html>

## 2. Bonnefon, J.-F., Shariff, A., & Rahwan, I. The social dilemma of autonomous vehicles. <sup>2</sup>

В статье содержится подробный разбор этических проблем выбора на основе нескольких опросов, ряд из которых посвящен именно закреплению на программном уровне в качестве одной из основ беспилотных автомобилей проблемы «вагонетки». Результаты исследований показали, что люди склонны поддержать внедрение технологий, которые будут работать на основе утилитарного подхода, однако себе приобрели бы такое беспилотное авто, которое бы стремилось спасти их любой ценой.

” Морального выбора нет. Всё объекты одинаково важны и все жизни ценны. Необходимо сделать все чтобы смягчить последствия столкновения.

КамАЗ

## 3. Himmelreich, J. Never Mind the Trolley: The Ethics of Autonomous Vehicles in Mundane Situations. <sup>3</sup>

Автор выдвигает предположение, что проблема вагонетки никак не соотносится с реальными ситуациями, возникающими при вождении.

Он пишет о четырех проблемах:

1. Проблема «вагонетки», учитывая технические ограничения, основана на предположениях, которые противоречат друг другу.
2. Проблема «вагонетки» освещает лишь ограниченный круг этических вопросов в той мере, в какой они согласуются с определенными рамками.
3. Проблема «вагонетки» требует морального ответа, когда требуется политический ответ.
4. Решение проблемы «вагонетки» может быть эпистемологически проблематичным (например, презюмируется, что все вовлеченные лица идентичны по своим личностным характеристикам, что в жизни зачастую не так, и т.д.).

---

2. Bonnefon, J.-F., Shariff, A., & Rahwan, I. (2016). The social dilemma of autonomous vehicles . Science, 352(6293), 1573–1576 // URL: doi:10.1126/science.aaf2654

3. Himmelreich, J. Never Mind the Trolley: The Ethics of Autonomous Vehicles in Mundane Situations. Ethic Theory Moral Prac 21, 669–684 (2018) // URL: doi:10.1007/s10677-018-9896-4

#### 4. **Bryant W. Smith. The Trolley and the Pinto: Cost-Benefit Analysis in Automated Driving and Other CyberPhysical Systems.**<sup>4</sup>

Автор этой статьи уверен, что сосредоточение внимания общественности на проблеме «вагонетки» отвлекает от гораздо более насущных этических проблем. Даже если автоматизированные системы в итоге окажутся более способными, чем люди, они не будут ни всеведущими, ни всемогущими: автоматизация не означает прекращение неопределенности. Предполагая такое всеведение, проблема «вагонетки» бесполезно сужает дискуссию от риска до простого вреда.

” В мире не разработано ни одной системы беспилотных транспортных средств, которая была бы по-настоящему доверенной и безопасной. Пройдет не один год до того, как система обучится на реальных, а не синтетических данных; должна быть перестроена инфраструктура и адаптирована городская среда, чтобы проблем безопасности не возникало.

Илья Попов, ИТМО

#### 5. **B. Lundgren. Safety requirements vs. crashing ethically: what matters most for policies on autonomous vehicles.**<sup>5</sup>

Автор считает, что существуют серьезные методологические проблемы, связанные с тем, как в настоящее время проводится дискуссия по этической проблеме «вагонетки» как в философско-этической литературе, так и в эмпирической литературе, и в публичных дебатах. Кроме того, дискуссия связана с и является второстепенной по отношению к более важным вопросам требований безопасности и аргументации в пользу безопасности. В статье изучается и даются краткие выводы по большому количеству работ, оспаривающих важность проблемы вагонетки.

” В настоящий момент полагаем преждевременным доверять алгоритмам искусственного интеллекта, применяемым в технологиях автономного вождения транспортных средств, право оценки рисков причинения вреда и выбора вариантов действий для минимизации последствий.

SberAutoTech

- 
4. Bryant Walker Smith, The Trolley and the Pinto: Cost-Benefit Analysis in Automated Driving and Other Cyber-Physical Systems, 4 Tex. A&M L. Rev. 197 (2017) // URL: <https://scholarship.law.tamu.edu/lawreview/vol4/iss2/5/>
  5. Lundgren, B. Safety requirements vs. crashing ethically: what matters most for policies on autonomous vehicles. AI & Soc 36, 405–415 (2021) // URL: <https://doi.org/10.1007/s00146-020-00964-6>

## 6. Noah J. Goodall. **Away from Trolley Problems and Toward Risk Management.**<sup>6</sup>

Хотя проблема «вагонетки» представляет собой интересную и морально неоднозначную ситуацию, она не является основной этической проблемой для высокоавтоматизированных ТС – к такой мысли приходит автор данной статьи. Транспортные средства должны постоянно оценивать опасность действий и сравнивать ценность различных объектов на дороге. То, как формулируется ценность, и как риски, связанные с вождением, распределяются между другими участниками дорожного движения, является большей этической проблемой.

” Если столкновение беспилотного автомобиля неизбежно, стоит просчитывать и выбирать сценарий с максимальной вероятностью безопасного исхода с постоянным обновлением информации о дорожной обстановке и наиболее безопасных сценариях в текущий момент.

И.Дейлид, Иннополис

## 7. A. D’Amato et al., **Exceptional Driving Principles for Autonomous Vehicles.**<sup>7</sup>

В работе на основании анализа этической проблемы «вагонетки», существующего правового регулирования и технических особенностей высокоавтоматизированных ТС выводится ряд принципов, в частности: (1) система автопилотирования должна быть запрограммирована таким образом, чтобы стремиться к обеспечению безопасности каждого участника дорожного движения; и (2) система автопилотирования должна быть таким образом запрограммирована, чтобы не соблюдать конкретные положения правил дорожного движения, если это необходимо для смягчения или предотвращения столкновения.

” Если столкновение неизбежно, система должна сделать все, чтобы смягчить последствия такого столкновения.

КамАЗ

---

6. Noah J. Goodall (2016). Away from Trolley Problems and Toward Risk Management. Applied Artificial Intelligence, 30(8):810-821 // URL: doi:10.1080/08839514.2016.1229922

7. A. D’Amato et al., Exceptional Driving Principles for Autonomous Vehicles, 2022 J. L. & Mob. 2. // URL: <https://repository.law.umich.edu/ilm/vol2022/iss1/2/>

## 8. Похожую позицию с Н. Гудэллом и А. Д’Амаро занимает менеджер компании Mercedes-Benz по системам помощи водителю и активной безопасности Кристоф фон Хьюго:<sup>8</sup>

«Мы считаем, что эта этическая проблема не так актуальна, как люди сегодня думают. Это ситуация, с которой нынешние водители не могут справиться, и с физической точки зрения мы не можем предотвратить ее сегодня, как не могут и высокоавтоматизированные транспортные средства (ВАТС). Просто ВАТС намного лучше, чем среднестатистические водители-люди. Это моральный вопрос: кого мы спасаем? Девяносто девять процентов нашей инженерной работы заключается в том, чтобы убедиться, что такой ситуации вообще не возникнет.

” На данный момент проблема является скорее теоретической. Так как обычно, в подобных ситуациях автопилот старается в первую очередь затормозить, а не выбрать, кого сбить».  
Дмитрий Куликов, технический директор ООО «Эвокарго»

Мы упорно работаем над тем, чтобы ВАТС не попадали в подобные ситуации и чтобы они держались подальше от потенциальных ситуаций, в которых им пришлось бы принимать подобные решения.»

## 9. Аналогичное мнение высказал и Эндрю Чатем, главный инженер проекта компании X, аффилированной с Google, занимающейся разработкой высокоавтоматизированных транспортных средств.<sup>9</sup>

«Этот вопрос не так интересен, как кажется. Во всех поездках мы никогда не оказывались в ситуации, когда приходилось выбирать между детской коляской и бабушкой. Даже если бы вы действительно оказались в подобной ситуации, обычно это означало бы, что вы допустили ошибку на пару секунд раньше. И поэтому, как высоконравственный инженер-программист, приходящий на работу в офис, если я хочу спасти жизни, моя цель - не допустить, чтобы мы попали в такую ситуацию, потому что это означает, что мы облажались. Это лишает проблему некоторой интеллектуальной интриги, но ответом почти всегда является “нажать на тормоза”.»

” Наш алгоритм распознавания не делает различий между людьми разного возраста или демографической группы. Он реагирует только в зависимости от дистанции, скорости, протяженности препятствий. Также мы учитываем потенциальное воздействие на препятствие, если транспортное средство столкнется с ним. Следовательно, этические соображения пока не являются основным решающим фактором в поведении транспортных средств.

BAIDU

8. A. D’Amato et al., Exceptional Driving Principles for Autonomous Vehicles, 2022 J. L. & Mob. 2. // URL: <https://repository.law.umich.edu/jlm/vol2022/iss1/2/>

9. Self-Driving Mercedes Will Be Programmed To Sacrifice Pedestrians To Save The Driver // URL: <https://www.fastcompany.com/3064539/self-driving-mercedes-will-be-programmed-to-sacrifice-pedestrians-to-save-the-driver>

## 10. Отдельного внимания заслуживает эксперимент Moral Machine.<sup>10</sup>

Исследователи из MIT Media Lab создали платформу «Moral Machine» (<https://www.moralmachine.net/>), на которой людям предлагали сделать выбор по проблеме вагонетки в разных интерпретациях. То есть, эта работа посвящена не важности самой проблемы или ее практической ценности, а скорее отношению обычных людей к потенциальному «поведению» высокоавтоматизированных ТС, загнанных в угол проблемой «вагонетки».

**” Обсуждение возможности доверия алгоритмам искусственного интеллекта оценивать риски и выбирать варианты действий для минимизации последствий при управлении транспортными средствами целесообразно после накопления соответствующего опыта эксплуатации полностью автономных транспортных средств.**

В рамках эксперимента проверяли девять факторов: спасение людей (по сравнению с домашними животными), сохранение курса (по сравнению с отклонением от курса), сохранение пассажиров (по сравнению с пешеходами), сохранение большего количества жизней (по сравнению с меньшим количеством жизней), сохранение мужчин (по сравнению с женщинами), сохранение молодых (по сравнению с пожилые люди), спасение пешеходов, которые переходят дорогу на законных основаниях (по сравнению с переходом в неположенном месте или на красный сигнал светофора), спасение тех, кто в хорошей физической форме (по сравнению с более полными), и спасение тех, кто имеет более высокий социальный статус (по сравнению с низким социальным статусом). Некоторые персонажи обладали другими атрибутами (такими как беременность, врачи, преступники и т. д.), которые не попадали в эти проверяемые факторы. Результаты сформированы из более 40 миллионов ответов миллионов пользователей из 233 стран мира. Во всем мире участники отдавали предпочтение человеческим жизням, а не жизням животных, таких как собаки и кошки. Они хотели сохранить больше жизней, чем меньше, а также хотели сохранить более молодые жизни по сравнению со старыми. Чаще всего щадили младенцев, меньше всего щадили кошек. Что касается гендерных различий, врачи-мужчины и старики были спасены чаще, чем врачи-женщины и пожилые дамы. В то время как спортсменов и более крупных женщин щадили больше, чем спортсменов-мужчин и более крупных мужчин. Большинство также предпочитали щадить пешеходов, а не пассажиров, и законопослушных лиц, а не правонарушителей.

---

10. Self-Driving Mercedes Will Be Programmed To Sacrifice Pedestrians To Save The Driver // URL: <https://www.fastcompany.com/3064539/self-driving-mercedes-will-be-programmed-to-sacrifice-pedestrians-to-save-the-driver>