

В.А. Новицкий, В.В. Паршонков
ОШИБКА ИСКУССТВЕННОГО ИНТЕЛЛЕКТА
КАК ОБСТОЯТЕЛЬСТВО НЕПРЕОДОЛИМОЙ СИЛЫ
(ФОРС-МАЖОР)

Виталий Анатольевич Новицкий – доцент кафедры гражданского права и процесса, Санкт-Петербургский университет технологий управления и экономики, кандидат юридических наук, доцент, адвокат АК «Правовой Петербург», г. Санкт-Петербург; **e-mail: 78jurist@gmail.com**.

Виталий Викторович Паршонков – аспирант кафедры гражданского права и процесса, Санкт-Петербургский университет технологий управления и экономики, г. Санкт-Петербург; **e-mail: vitaliyvictorovichlaw@gmail.com**.

Анализируется проблема квалификации ошибок искусственного интеллекта в качестве обстоятельства непреодолимой силы (форс-мажора). Авторами проведено комплексное исследование доктринальных положений о регулировании ответственности искусственного интеллекта. Результаты исследования свидетельствуют о наличии множества подходов, которые регулируют ответственность искусственного интеллекта по-разному. Сформулированы случаи, в которых ошибка искусственного интеллекта может быть квалифицирована в качестве обстоятельства непреодолимой силы (форс-мажора).

Ключевые слова: искусственный интеллект; обстоятельство непреодолимой силы; форс-мажор; ответственность; ошибки искусственного интеллекта; обязательства.

V.A. Novitsky, V.V. Parshonkov
ARTIFICIAL INTELLIGENCE ERROR AS FORCE MAJEURE

Vitaly Novitsky – senior lecturer, the Department of Civil Law and Process, St. Petersburg University of Management Technologies and Economics, PhD in Law, associate professor, lawyer, the law office “Pravovoy Peterburg”, St. Petersburg; **e-mail: 78jurist@gmail.com**.

Vitaly Parshonkov – post-graduate student, the Department of Civil Law and Process, St. Petersburg University of Management Technologies and Economics, St. Petersburg; **e-mail: vitaliyvictorovichlaw@gmail.com**.

We analyze the problem of qualifying artificial intelligence errors as force majeure. We carry out a comprehensive study of the doctrinal provisions concerning the regulation of artificial intelligence liability. As a result, we have discovered a variety of approaches that regulate the responsibility of artificial intelligence in different ways. We specify certain cases when an artificial intelligence error can be qualified as force majeure.

Keywords: artificial intelligence; force majeure; liability; artificial intelligence errors; obligations.

Искусственный интеллект (далее – ИИ) разработан с целью улучшения эффективности и точности решения сложных задач. Несмотря на многочисленные преимущества, ИИ не является идеальным инструментом и может допускать ошибки. Возникает вопрос с позиций права о том,

могут ли быть ошибки ИИ обстоятельством непреодолимой силы.

Ошибка ИИ может быть определена как некорректное решение или предсказание, полученное с помощью алгоритмов цифрового (машинного) обучения. Полагаем, что ошибки ИИ объективно приводят

к негативным последствиям. Такая ситуация может быть определена в правоведении как обстоятельство непреодолимой силы. Совершение ошибок ИИ неизбежно, и иногда их невозможно предотвратить. Они, например, могут быть вызваны работой с недостаточным количеством данных либо обусловлены ситуациями, в которых алгоритм не может учесть все переменные в сложных случаях.

Ошибки ИИ могут иметь существенные последствия в различных областях жизни. Например, в медицине ИИ используется для диагностики болезней и принятия решений о лечении [9]. Если алгоритм совершает ошибку, это может привести к неверному диагнозу или неверному лечению. В автомобильной промышленности следствием ошибок ИИ могут стать аварии и даже смертельный исход [21]. Например, если автоматическая система торможения в автомобиле не работает должным образом, это может привести к столкновению и травмам для водителя и пассажиров. В результате ошибок ИИ могут возникнуть социальные проблемы. Например, если алгоритм машинного обучения используется для принятия решений о найме сотрудников, это может привести к дискриминации по расовым, половым или иным признакам [10].

На первый взгляд, если автономная система начинает работать со сбоями, оператор должен нести ответственность, как и в случае любого другого инструмента, оператором которого выступает человек. Если возможные неисправности искусственного интеллекта предсказуемы для его разработчика, то это может говорить о возможности применения ответственности в его отношении.

Сегодня в юридической доктрине существуют разные подходы в отношении гражданско-правовой ответственности за вред, причиненный искусственным интеллектом [6]:

1. Модель, в рамках которой система ИИ рассматривается как принципиально невинный агент, инструмент реального исполнителя правонарушения, а противоправное деяние совершает человек. Им может быть разработчик, собственник (вла-

делец) и пользователь таких систем (далее – модель ответственности № 1).

2. Модель естественных вероятностных последствий, которая предполагает, что система ИИ реализует действия, являющиеся естественным и вероятным последствием программирования. В рамках нее предусмотрено, что лицо, запрограммировавшее машину (разработчик), проявило преступную небрежность или самонадеянность (далее – модель ответственности № 2).

3. Модель прямой ответственности системы ИИ за свои действия. Она не получает широкого распространения в современных условиях ввиду того, что, несмотря на автономность или обучаемость ИИ, он все же не обладает человеческим сознанием и его правосубъектностью и не может быть привлечен к ответственности (далее – модель ответственности № 3).

Из приведенных подходов следует, что сегодня отсутствует единое мнение относительно вопроса об определении причинителя вреда при использовании ИИ. Кроме того, в настоящее время отсутствует и правовое регулирование данной проблемы, при этом подходы к способам нормативного регулирования ИИ могут быть различны и будут исследованы авторами в дальнейшем.

Модель ответственности № 2 уже находит отражение в судебной практике США. 21 апреля 2023 г. Высший суд округа Лос-Анджелес вынес решение по иску о возмещении ущерба в связи с тем, что во время движения автомобиль *Tesla*, управляемый автопилотом, допустил наезд на бордюр из-за совершения непредвиденного маневра.

В иске истец указывал, что автомобиль двигался со скоростью примерно от 25 до 30 миль в час, при этом автопилот не смог распознать центральную медиану. Внезапно автопилот дал сбой, и машина свернула в центральную полосу. Колесо автомобиля ударило о бордюр, в результате сработали подушки безопасности. Водитель не успел отреагировать, но она попыталась защитить лицо, убрав руки с руля и расположив их перед лицом [19].

Суд присяжных решил, что компания-

разработчик не несет ответственности в данном случае. Более того, присяжные утверждали, что *Tesla* четко предупредила, что частично автоматизированное программное обеспечение для вождения не является самоуправляемой системой. Причина аварии заключается еще и в том, что водитель отвлекся в пути [20].

Стоит отметить, что в ходе проведения исследования авторами не было обнаружено релевантной судебной практики отечественных судов, относимых к теме настоящего исследования.

Таким образом, первые судебные разбирательства по спорам в отношении ошибок ИИ складываются на основе применения модели ответственности № 2 и предполагают, что ответственность может быть возложена на владельца (пользователя) системы, управляемой ИИ, если им не совершены действия по предотвращению возникновения ущерба. При этом наступление ущерба было предотвратимым.

Следует определить, в каких случаях ошибка ИИ может быть отнесена к форс-мажору. В отечественном праве легальное определение непреодолимой силы (форс-мажора) закреплено в п. 3 ст. 401 Гражданского кодекса Российской Федерации (ГК РФ): это – чрезвычайные и непредотвратимые при конкретных условиях обстоятельства, делающие невозможным надлежащее исполнение обязательств [1].

Институт форс-мажоров призван защитить интерес стороны договора при наступлении обстоятельств, которые находятся вне ее контроля. В первую очередь следует определить, насколько контролируемы со стороны пользователя (владельца) или разработчика действия ИИ при его использовании. Напомним, под чрезвычайностью понимаются исключительные, необычные, выходящие за пределы «нормального» случаи, условия [2]. Под непредотвратимостью понимают невозможность любого участника гражданского оборота, осуществляющего аналогичную деятельность, избежать наступивших последствий.

Попытаемся раскрыть суть понятия «искусственный интеллект». В общем понимании этот термин является обобща-

ющим. Его используют при характеристике отрасли компьютерных наук, и он означает способность машины или программы приобретать и применять знания, а также действовать естественно, то есть так, как человек [8]. По своей сути ИИ работает на основе заложенного в него разработчиком алгоритма, позволяющего ему обучаться или быть обученным.

В настоящее время отсутствует общепринятое понятие ИИ [7]. Поэтому в статье предлагаем понимать ИИ в широком смысле, как это отражено в приведенном выше определении. С учетом ранее изложенных подходов об ответственности ИИ можно сформулировать следующие модели применения института форс-мажоров относительно ошибок ИИ:

1. Невозможность обращения к форс-мажору в случае ошибки ИИ в рамках модели ответственности № 1. В данном случае ошибка ИИ не будет являться форс-мажором, поскольку концепция предполагает, что ИИ служит лишь инструментом. Это означает, что форс-мажор должен наступить в момент совершения разработчиком, владельцем или пользователем системы действия, которое повлекло последствия в виде ошибки ИИ.

Указанная концепция также находит отражение в п. 9.1.1 Рекомендации Парламентской ассамблеи Совета Европы от 28 апреля 2017 г. № 2102 «Слияние с технологиями, искусственный интеллект и права человека» [14]. Данный пункт гласит о том, что ответственность за действия ИИ всегда лежит на человеке, а ссылка на его независимые решения не может освободить от ответственности даже в случаях, если это действие не совершено по указанию человека. Аналогичная позиция следует из Резолюции Генеральной ассамблеи ООН от 12 декабря 2001 г. № 56/83 «Ответственность государств за международно-противоправные деяния» (подп. «а» ч. 2 ст. 23) [18].

2. Модель ответственности № 2 позволяет квалифицировать ошибку ИИ как форс-мажор. Согласно этой концепции, система ИИ реализует действия, являющиеся естественным и вероятным последствием программирования. Данная концепция

предполагает, что именно разработчик ИИ должен обеспечить недопущение вредоносных действий [17]. Однако, если машинное обучение проходит не изолированно, а посредством самообучения в интернете или взаимодействия с неопределенным кругом лиц, то гарантировать недопущение вредоносных действий становится затруднительным.

Так, известен случай, который произошел с чат-ботом Тэй, запущенным компанией *Microsoft* и обучавшимся посредством общения с аудиторией американских мессенджеров и социальных сетей. В результате общения с аудиторией, несмотря на предусмотренную разработчиками защиту, чат-бот начал активно оскорблять некоторые социальные группы людей [11]. Иными словами, злоумышленники могут целенаправленно генерировать вредоносную информацию для провоцирования ИИ на ошибки. Даже контроль над обучением ИИ со стороны человека не гарантирует отсутствие ошибок, поскольку такой контроль, как правило, осуществляется на основе контрольной выборки, имеющей объективные границы.

В связи с изложенным необходимо ответить на вопрос о том, является ли ошибка, допущенная ИИ на основе полученной в ходе обучения вредоносной информации, чрезвычайной и непредотвратимой. Ответ на него зависит от того, какой подход к пониманию форс-мажоров будет применен – субъективный или объективный.

Согласно объективной теории форс-мажоров событие, которое повлекло причинение вреда, является исключительным, его невозможно предотвратить. Данное событие характеризуется тем, что оно возникает извне, не связано с деятельностью ответственного лица [3]. Субъективная теория предполагает, что форс-мажор – это событие, которое невозможно предотвратить никакими мерами. В отличие от объективной теории, для квалификации события используется как признак исключительности, так и непредотвратимости.

Например, О.С. Иоффе определял непреодолимую силу (форс-мажор) как внешнее или внутреннее по отношению к

вредоносной деятельности событие, которое, будучи чрезвычайным по своему характеру, непредотвратимо с применением хозяйственно доступных для такого лица средств [5, с. 153]. Рассуждая о непредотвратимости события доступными для лица средствами, нельзя не обратить внимание на то, что при обучении ИИ посредством общения с большой аудиторией разработчик не может контролировать в целом поток данных на предмет их вредоносности ввиду огромного объема последних.

Таким образом, обращение к субъективной теории дает возможность квалифицировать ошибку ИИ как форс-мажор, если разработчик сделал все, что от него зависело для обеспечения нормального функционирования ИИ, но ввиду внешних обстоятельств, принятые разработчиком меры не смогли предотвратить возникновение вредных последствий. При таких обстоятельствах ошибку ИИ можно квалифицировать как чрезвычайное обстоятельство, поскольку ИИ по своей сути должен действовать в рамках, предписанных для него правил, а отклонение от них не может быть обычным и не выходящим за пределы его нормального функционирования ввиду заранее определенных разработчиком рамок деятельности. Ошибка будет соответствовать признаку непредотвратимости, поскольку полный контроль за обрабатываемыми ИИ данными объективно невозможен для ограниченного штата сотрудников разработчика.

3. Может ли ошибка ИИ рассматриваться в качестве форс-мажора при использовании модели прямой ответственности системы ИИ за свои действия (модель ответственности № 3)?

Для обоснования возможности или невозможности квалификации ошибки ИИ в качестве форс-мажора в рамках исследуемой модели сделаем несколько обоснованных дополнений. В частности, ИИ может быть наделен законодателем правосубъектностью, и в данном случае он станет нести ответственность за свою ошибку, которая, по нашему мнению, не будет признана форс-мажором. Однако это видится возможным лишь в отношении «сильного» ИИ [16], способного решать любые задачи

подобно человеческому мозгу.

Было бы сомнительным относить, например, автопилот машины к ИИ, который должен быть наделен правосубъектностью, так как он решает лишь строго ограниченный круг задач по пилотированию автомобиля. На данном этапе ИИ не может быть наделен правосубъектностью и нести ответственность за свои действия. При этом система ИИ не имеет таких характеристик, как сознание, представление о морали, осознание окружающей действительности, позволяющих людям понимать, совершают ли они противоправные действия или не совершают их.

В этой связи существует проблема, которая заключается в отсутствии понимания возможности применения существующих правовых категорий к автономным ИИ [13]. Приведем известный пример о продвинутом ИИ. Речь идет о *ChatGPT*, разработанном американской компанией *OpenAI*. Последняя на момент написания настоящей статьи версия указанного чат-бота смогла обмануть человека для выполнения поставленной ему задачи [12].

Данный пример показывает, что чем более совершенным является ИИ, тем более непредсказуемым и свободным в своих действиях он становится. Так, Питер Асаро считает, что ИИ может быть наделен некоторой правосубъектностью. Он полагает, что самообучение может обуславливать совершение ИИ заранее не запрограммированных действий [15]. Например, в п. «z» Резолюции Европарламента «Нормы гражданского права о робототехнике» от 16 февраля 2017 г. указано, что роботы способны «осуществлять когнитивные процессы (учиться на собственном опыте и принимать квазинезависимые решения), роботы все больше и больше становятся похожи на агентов, которые могут взаимодействовать со своей средой и вносить в нее изменения» [13].

Если мы допускаем возможность надления ИИ правосубъектностью или квазиправосубъектностью, то при каких условиях можно утверждать, что его ошибка является форс-мажором? Б.С. Антимонов полагал, что поведение другого лица при некоторых условиях может рассматривать-

ся как действие непреодолимой силы, то есть «поведение другого лица должно быть объективно-случайным для ответственного лица» [4, с. 190–191].

Данное условие, видимо, применимо к человеку, поскольку он обладает собственной волей, а в форс-мажоре не должно быть волевого момента. Однако применимо ли понятие «воля» к ИИ в том смысле, в котором оно применяется к человеку? Ответ на этот вопрос не найден, а значит, он остается открытым.

С учетом ранее изложенных доводов при описании возможности квалификации ошибки ИИ в качестве форс-мажора можно резюмировать.

1. В рамках модели принципиально невинного агента невозможно квалифицировать ошибку ИИ в качестве форс-мажора, так как данная концепция предполагает, что разработчик, собственник (владелец) и пользователь несут ответственность за действия ИИ в любом случае.

2. Модель естественных вероятностных последствий допускает применение форс-мажора, поскольку в рамках такой модели ошибка ИИ будет соответствовать критериям для ее квалификации в качестве такового.

3. В результате исследования возможности квалификации ошибки ИИ в качестве форс-мажора сделан вывод об отсутствии достаточного для этих целей категориального аппарата. При таких обстоятельствах невозможно сформулировать хотя бы сколько-нибудь верную гипотезу в контексте исследования модели прямой ответственности ИИ.

Из вышеизложенного следует, что порядок применения форс-мажора должен различаться как в зависимости от модели ответственности ИИ, так и в зависимости от типа ИИ: чем автономнее ИИ, тем сложнее определить правовой статус его действий в рамках существующего в настоящее время законодательства.

Однако в ряде случаев ошибка ИИ может быть рассмотрена как обстоятельство непреодолимой силы (форс-мажор), поскольку при известных условиях ошибка ИИ может быть квалифицирована как чрезвычайная, неизбежная и непредотвращаемая.

тимая. При отсутствии единого подхода к возникновению ответственности в отношении ИИ вопрос о квалификации его ошибки в качестве форс-мажора фактически можно считать вопросом политики права, он зависит от принимаемого тем или иным правопорядком режима ответственности.

В итоге целесообразно заключить, что для выявления и определения наиболее эффективных путей регулирования исследуемого вопроса необходимо и в дальнейшем продолжить тщательное изучение отраженных в настоящей статье проблем и способов их решения, в том числе исследование концептуально иных подходов к квалификации ошибки ИИ в качестве форс-мажора.

ЛИТЕРАТУРА

1. Гражданский кодекс Российской Федерации (часть первая): Федер. закон от 30 ноября 1994 г. № 51-ФЗ (в ред. от 05.12.2022 г.) // Российская газета. 1994. 8 декабря.
2. О применении судами некоторых положений Гражданского кодекса Российской Федерации об ответственности за нарушение обязательств: Постановление Пленума Верховного Суда РФ от 24 марта 2016 г. № 7 // Справ.-правовая система «КонсультантПлюс». URL: https://www.consultant.ru/document/cons_doc_LAW_195783/137415694d421422854b97e7ae908c8d9f31546b/?ysclid=1ifzkpm832817626816 (дата обращения: 17.05.2023).
3. Советское гражданское право / под ред. С.И. Вильнянского. М.: Юриздат, 1940. 246 с.
4. Антимонов Б.С. Гражданская ответственность за вред, причиненный источником повышенной опасности. М.: Госюриздат, 1952. 296 с.
5. Иоффе О.С. Обязательственное право. М.: Юрид. лит., 1975. 880 с.
6. Карпенко О.А. Ответственность искусственного интеллекта в современной правовой системе // Государственное регулирование общественных отношений в регионе: социально-экономические, правовые и историко-культурные аспекты: сборник науч. ст. / отв. ред. Е.А. Ильина, Г.Л. Белов. Чебоксары: Чуваш. гос. пед. ун-т, 2022. С. 213–214.
7. Морхат П.М. К вопросу об определении понятия искусственного интеллекта // Право и государство: теория и практика. 2017. № 12 (156). С. 25–32.
8. Терехина Е.А., Селюкова Г.П. Об искусственном интеллекте // Актуальные вопросы науки и хозяйства: новые вызовы и решения: сборник материалов V студенч. науч.-практ. конф., Тюмень, 17–19 марта 2021 г. Ч. 2. Тюмень: Государственный аграрный университет Северного Зауралья, 2021. С. 581–584.
9. Фершт В.М., Латкин А.П., Иванова В.Н. Современные подходы к использованию искусственного интеллекта в медицине // Территория новых возможностей. 2020. Т. 12. № 1 (48). С. 121–130.
10. Щербакова О.В. Использование программ искусственного интеллекта при найме работников // Электронное приложение к Российскому юридическому журналу. 2021. № 3. С. 72–76.
11. Почему девушка-чатбот от Microsoft поддалась влиянию троллей с 4chan // Хабр. 2016. 25 марта. URL: <https://habr.com/ru/articles/392129> (дата обращения: 20.05.2023).
12. ChatGPT обманул человека, чтобы решить поставленную задачу // Ixbt. 2023. 17 марта. URL: <https://www.ixbt.com/news/2023/03/17/chatgpt-obmanul-cheloveka-chtoby-reshit-postavlennuju-zadachu.html> (дата обращения: 20.05.2023).
13. European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL)) // EUR-Lex. URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52017IP0051&from=EN> (дата обращения: 06.05.2023).
14. Technological convergence, artificial intelligence and human rights: Recommendation № 2102 (2017) of Parliamentary Assembly of the Council of Europe on 28 April 2017 // Parliamentary Assembly. URL: <http://assembly.coe.int/nw/xml/XRef/Xref-XML2HTML-en.asp?fileid=23726&lang=en> (дата обращения: 17.05.2023).

15. *Asaro P.M.* Robots and Responsibility from a Legal Perspective // Proceedings of the IEEE. 2007. Vol. 4. № 14. P. 20–24. URL: <https://peterasaro.org/writing/ASARO%20Legal%20Perspective.pdf> (дата обращения: 06.05.2023).
16. *Castro D., New J.* The Promise of Artificial Intelligence// Center for data innovation. 2016. 44 p. URL: <https://www2.datainnovation.org/2016-promise-of-ai.pdf> (дата обращения: 06.05.2023).
17. *Prakken H.* On how AI & law can help autonomous systems obey the law: a position paper // AI4J- Artificial Intelligence for Justice. 2016. P.42–46.
18. Articles on Responsibility of States for Internationally Wrongful Acts // Audiovisual Library of International Law. URL: https://www.un.org/ru/documents/decl_conv/conventions/pdf/responsibility.pdf (дата обращения: 20.05.2023).
19. Complaint to the superior court of the state of California for the county of Los Angeles // Arstechnica. 2023. URL: <https://cdn.arstechnica.net/wp-content/uploads/2023/04/tesla-justine-hsu-complaint.pdf> (дата обращения: 17.05.2023).
20. *Roy A., Levine D., Jin H.* Tesla wins bellwether trial over Autopilot car crash // Reuters.2023. April 22. URL: <https://www.reuters.com/legal/us-jury-set-decide-test-case-tesla-autopilot-crash-2023-04-21> (дата обращения: 17.05.2023).
21. *Siddiqui F., Lerman R., Merrill J.B.* Teslas running Autopilot involved in 273 crashes reported since last year // The Washington Post. 2022. June 15. URL: <https://www.washingtonpost.com/technology/2022/06/15/tesla-autopilot-crashes/> (дата обращения: 07.05.2023).