
ИНФОРМАЦИОННОЕ ПРАВО
И ИНФОРМАЦИОННАЯ БЕЗОПАСНОСТЬ

УДК 341.9

КАТЕГОРИЯ «ЮРИДИЧЕСКАЯ ОТВЕТСТВЕННОСТЬ»
В КОНТЕКСТЕ ЕЕ ПРИМЕНЕНИЯ К СИСТЕМАМ
ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

© 2024 г. С. С. Зенин*, И. М. Япринцев**, Л. В. Иванова***

Тюменский государственный университет

*E-mail: s.s.zenin@utmn.ru

**E-mail: i.m.yapryntsev@utmn.ru

***E-mail: l.v.ivanova@utmn.ru

Поступила в редакцию 10.09.2024 г.

Аннотация. В статье рассматривается категория «юридическая ответственность» с точки зрения отечественной и зарубежной юридической науки, а также оценивается возможность ее применения к характеристикам искусственного интеллекта как автономного волевого субъекта и возможности его наделения деликтоспособностью. Исследователи практически не рассматривают искусственный интеллект как самостоятельный деликтоспособный субъект права, притом что прогнозы развития данной сферы общественных отношений на ближайшие десятилетия (в частности, создание и внедрение сильного искусственного интеллекта) говорят о необходимости рассмотрения такой категории и ее потенциального юридического закрепления.

Ключевые слова: искусственный интеллект, право, цифровые технологии, юридическая ответственность, правонарушение, санкция, наказание, информация, информационное общество.

Цитирование: Зенин С.С., Япринцев И.М., Иванова Л.В. Категория «юридическая ответственность» в контексте ее применения к системам искусственного интеллекта // Государство и право. 2024. № 11. С. 163–173.

Исследование выполнено за счет гранта Российского научного фонда № 24-28-01112, <https://rscf.ru/project/24-28-01112/>

DOI: 10.31857/S1026945224110145

CATEGORY OF “LEGAL RESPONSIBILITY” IN THE CONTEXT
OF ITS APPLICATION TO ARTIFICIAL INTELLIGENCE SYSTEMS

© 2024 S. S. Zenin*, I. M. Yapryntsev**, L. V. Ivanova***

Institute of State and Law of Tyumen State University

*E-mail: s.s.zenin@utmn.ru

**E-mail: i.m.yapryntsev@utmn.ru

***E-mail: l.v.ivanova@utmn.ru

Received 10.09.2024

Abstract. The article examines the category of “legal responsibility” from the point of view of domestic and foreign legal science, and also assesses the possibility of its application to the characteristics of artificial intelligence as an autonomous volitional subject and the possibility of endowing it with delinquency. Researchers practically do not touch on artificial intelligence as an independent delinquent subject of law, despite the fact that forecasts of the development of this sphere of public relations for the coming decades (in particular, the creation and implementation of strong artificial intelligence) indicate the need to consider such a category and its potential legal consolidation.

Key words: artificial intelligence, law, digital technologies, legal liability, offence, sanction, punishment, information, information society.

For citation: Zenin, S.S., Yapryntsev, I.M., Ivanova, L.V. (2024). Category of “legal responsibility” in the context of its application to artificial intelligence systems // Gosudarstvo i pravo=State and Law, No. 11, pp. 163–173.

The research was carried out at the expense of a grant from the Russian Science Foundation No. 24-28-01112,
<https://rscf.ru/project/24-28-01112/>

Юридическая ответственность является неотъемлемой частью правовой культуры и юридической практики всех государств на всех этапах их исторического развития. В научных исследованиях, а также в правоприменительной практике сложилось относительно устойчивое понимание этого института, однако отдельные его аспекты продолжают исследоваться и оспариваться. Появление и активное внедрение в современную жизнь технологий искусственного интеллекта вывело данный вопрос на совершенно новый уровень, поскольку современные системы искусственного интеллекта представляют собой уникальную технологию, способную не только улучшать и развивать отдельные сферы человеческой деятельности, но и менять привычную картину современного общества в целом. Если рассматривать искусственный интеллект сквозь призму правового поля, очевиден процесс отставания развития общественных отношений и механизма их регулирования от технологических планок и соответствующих социальных вызовов, которые диктует в своем развитии искусственный интеллект. Одной из таких критически важных правовых проблем является юридическая ответственность, представляющая собой устоявшийся во времени и на практике механизм обеспечения правопорядка и социальной справедливости в русле закона и правовых ценностей общества.

В юридической научной литературе нет единого понимания юридической ответственности: она интерпретируется через понятие ответственности в общем, философском и социальном смыслах. В то же время, рассматривая различные точки зрения, можно сложить единую картину представления о юридической ответственности как эффективном инструменте регулирования общественных отношений. Как правило, ответственность предполагает элемент правосубъектности. Однако этот вопрос существенно осложняется новым для права понятием «искусственный интеллект».

Так, в частности, с точки зрения Г. Кельзена, автора «Чистого учения о праве», ответственность предусматривает обязанность индивида «вести себя определенным образом, если его противоположное поведение обуславливает акт принуждения в качестве санкции как меры воздействия»¹.

Г. Кельзен выделял: а) уголовно-правовую и гражданско-правовую ответственность; б) ответственность индивидуальную и коллективную; в) виновную ответственность (за преднамеренное деяние и за неосторожность); г) ответственность независимо от вины².

Г.Л. Харт рассматривает юридическую ответственность как отношение между возможным действием субъекта и представлением им последующего развития события и потенциальным действием субъекта, который тот должен или готов осуществить в целях предотвращения нежелательных вариантов развития событий и имеет место, когда мотивация одного из них обращена в будущее, хотя это и не означает, что она не может распространяться и на настоящее и прошлое (проспективная юридическая ответственность).

Ретроспективная же, по мнению Харта, обращена в прошлое как отношение между потенциальным действием субъекта, который тот должен или готов осуществить в целях предотвращения нежелательных вариантов развития событий, и уже наступившим событием от поступка³.

¹ Чистое учение о праве Ганса Кельзена // К XIII конгрессу Международной ассоциации правовой и социальной философии (Токио, 1987): сб. переводов / пер. С. В. Лезова; отв. ред. В. Н. Кудрявцев, Н. Н. Разумович. М., 1987. Вып. 1. С. 161.

² См.: там же. С. 164–166.

³ См.: Hart H. L.A. The Ascription of Responsibility and Rights // Proceedings of the Aristotelian Society. 1949. Vol. 49. Pp. 171–194 (см.: Харт Г.Л.А. Приписывание ответственности и прав / пер. с англ. В. В. Оглезнева // Правоведение. 2010. № 5 (292). С. 116–135).

Обобщая западноевропейский подход к теории юридической ответственности, в литературе представлена следующая классификация ее видов по различным основаниям⁴: темпоральный, показывающий время совершения действия с формулированием мотивации для его совершения или его легитимацией; прагматический, указывающий на основание возникновения отношения ответственности; субъектный, специфицирующий границы универсализации отношения ответственности; агентный, определяющий основной коррелят мотивации агента к совершению поступка⁵.

В отечественной юридической литературе признаки и принципы юридической ответственности расширяются и конкретизируются, однако на сегодняшний день их также недостаточно для полноценного их применения по отношению к действиям ИИ.

В. С. Нерсесянц понимает юридическую ответственность как меру правового принуждения за правонарушение, предусмотренную санкцией нарушенной нормы, и применяемую к правонарушителю компетентным государственным органом или должностным лицом в надлежащем процессуально-правовом порядке⁶.

Н. В. Витрук определяет юридическую ответственность как обратную сторону свободы. Автор подчеркивает, что, обладая избирательной активностью, определяя характер своего поведения, индивид постоянно должен соотносить его с социальными требованиями. Не подавляя личность, необходимость ставит ее в такое положение, когда она должна выбирать, решать самостоятельно. Необходимость выступает объективной предпосылкой ответственности, субъективной же предпосылкой служит свобода воли⁷.

В. М. Сырых среди признаков юридической ответственности в качестве ключевых выделяет следующие: это – вид государственного принуждения, реализуемого уполномоченными на то государственными органами (т.е. исходящий от государства); единственным основанием наступления юридической ответственности является

⁴ См.: Савельев Ю. М. О понятии юридической ответственности // Юрид. исследования. 2015. № 10. С. 61–80.

⁵ См.: Йонас Г. Принцип ответственности. М., 2004; Харт Г. Понятие права. СПб., 2003; Ingarden R. Über die Verantwortung. Ihre ontischen Fundamente. Stuttgart, 1970; Ingarden R. Kniążeczka o człowieku. Kraków, 1972.

⁶ См.: Нерсесянц В. С. Общая теория права и государства. М., 1999.

⁷ См.: Витрук Н. В. Общая теория юридической ответственности. М., 2008.

правонарушение; выражается в применении к правонарушителю негативной санкции⁸.

Такую же точку зрения поддерживает О. Э. Лейст, понимая правонарушение как отступление не от любой юридической обязанности, а только от той, которая носит императивный характер, снабжена юридической санкцией⁹ (в противовес иным нормонарушениям).

При рассмотрении положений о юридической ответственности важным является вопрос основания, при наличии которого ее применение допустимо. Следует согласиться, что такого рода основанием является юридический факт (либо их совокупность), представляющий собой правонарушение. Данная категория исследователями традиционно определяется как противоправное, виновное, наказуемое, общественно опасное деяние вменяемого лица, причиняющее вред интересам государства, общества и граждан¹⁰. Исходя из этого, правонарушение как правовая категория и как юридический факт имеет ряд признаков, а именно: противоправность; общественная опасность; виновность; интеллектуально-волевой характер; наказуемость.

Соответственно, верхнеуровневый¹¹ анализ подходов к пониманию такой категории, как «юридическая ответственность», в зарубежном и отечественном правопорядке позволяет выделить ряд ее признаков: закрепление в действующем законодательстве; наличие полного состава правонарушения; обеспеченность государственным принуждением; выражение в виде неблагоприятных для лица отрицательных последствий (лишение личных имущественных и неимущественных социальных благ); возложение и реализация в установленном законодательством порядке; наказание от имени государства уполномоченными компетентными органами и должностными лицами; выражение в виде санкций применительно к конкретному лицу в определенной ситуации (при этом не будучи тождественной самой санкции, которая может носить положительный характер, например, поощрение).

Наличие данных критериев и их социально-историческая устойчивость позволяет их соотнести с существующими подходами в контексте применения к системам искусственного интеллекта.

⁸ См.: Сырых В. М. Теория государства и права: учеб. М., 2012.

⁹ См.: Лейст О. Э. Санкции в советском праве. М., 1962. С. 47, 48.

¹⁰ См.: Матузов Н. И., Малько А. В. Теория государства и права. М., 2001. С. 432.

¹¹ Представляется, что сама по себе возможность распространения юридической ответственности на системы искусственного интеллекта не предполагает пересмотра фундаментальных основ данного юридического института.

С учетом обозначенных подходов к пониманию юридической ответственности, а также основываясь на уже имеющихся научных подходах к юридической ответственности в сфере искусственного интеллекта, допустимо выделение нескольких вариантов решения проблемы определения ответственного субъекта: ответственность создателя/разработчика искусственного интеллекта; ответственность конкретного пользователя или группы пользователей; ответственность самого искусственного интеллекта как субъекта права (в случае появления на научно-технологической арене «сильного» искусственного интеллекта).

Первые два подхода широко обсуждаются в научной сфере как в отечественной практике, так и за рубежом, и в своей основе связаны с определением того, кто или чьи действия являются первопричиной правонарушения с участием ИИ либо совершенного им самим, влекущего соответствующий вид юридической ответственности.

В качестве примера можно привести использование дипфейк-технологии для совершения мошеннических действий, когда присутствует конкретный деликтоспособный субъект, их совершающий, и установить личность которого возможно на практике. В этом случае ответственность лежит на пользователе, поскольку его как субъекта правонарушения можно идентифицировать (что упирается лишь в существующие технологические ограничения), а в его действиях есть прямой или косвенный умысел.

Ответственность разработчика может наступить в случае, если он преднамеренно либо в силу легкомыслия или небрежности оставил в разработанных алгоритмах существенные недоработки, которые привели к общественно опасным последствиям даже вследствие отсутствия противоправного умысла у пользователя (например, утечка персональных данных и нарушение медицинской тайны пациентов при использовании систем искусственного интеллекта для разработки медицинских рекомендаций). Сюда же можно отнести ошибки при диагностике заболеваний, следствием которых может быть причинение вреда здоровью или даже смерть пациента в случае, когда медицинский работник не имел возможности предотвратить такой вред и в его действиях/бездействии не было состава правонарушения.

Примером более сложного распределения юридической ответственности является использование алгоритмов искусственного интеллекта для выборки кандидатов при трудоустройстве на определенные должности. Российское трудовое законодательство запрещает дискриминацию по половому, расовому, национальному и иному признакам, однако искусственный интеллект в процессе выборки

потенциального кандидата может поступиться данной нормой и, например, ограничить выборку кандидатов для занятия определенной должности, убрав из нее женщин, представителей определенной нации или этноса, либо, наоборот, выбрав их в качестве предпочтительных кандидатур, тем самым отсекая потенциальную возможность трудоустройства другим кандидатам¹². В данном случае, установить ответственность не представляется возможным ни для одной из сторон, поскольку разработчик не закладывал подобного алгоритма и не имел умысла на дискриминацию тех или иных лиц, а работодатель, используя его продукт, не давал условий для подобной выборки кандидата.

Искусственный интеллект же, в свою очередь, в своем выборе руководствовался поиском в массиве данных с огромным количеством параметров, а также цифровых следов различных пользователей, включающих в себя господствующие стереотипы и массовые представления о представителях конкретной профессии, и, поскольку ограничить искусственный интеллект в данном случае затруднительно, его выборка носит дискриминирующий характер. Ответственность должна логически лежать на миллионах пользователей, оставлявших в сети информацию, на основании которой и были сделаны подобные выводы, поскольку именно их данные явились полем для его обучения и формирования выборки. Однако такого рода модель юридической ответственности не допустима.

Рассматривая вопрос об ответственности при использовании искусственного интеллекта, нельзя не упомянуть тот факт, что данное технологическое решение многими исследователями и практикующими специалистами позиционируется как источник повышенной опасности, концепция которого в этом контексте также рассматривается в работах отечественных и зарубежных ученых.

В частности, О.А. Ижаев, Д.Л. Кутейников, исследуя вопросы гражданско-правовой ответственности ИИ, предлагают деятельностный подход к определению источников повышенной опасности, суть которого изложена в постановлении Пленума Верховного Суда РФ от 26 января 2010 г. № 1, согласно которому по смыслу ст. 1079 ГК РФ источником повышенной опасности следует признавать любую деятельность, осуществление которой создает повышенную вероятность причинения вреда из-за невозможности полного контроля за ней со стороны человека, а также деятельность по использованию, транспортировке, хранению предметов, веществ и других объектов

¹² См.: Новиков Д.А. Трудовое право и искусственный интеллект: точки соприкосновения и расхождения // Ежегодник трудового права. 2024. № 14. С. 156–174.

производственного, хозяйственного или иного назначения, обладающих такими же свойствами¹³.

Авторы отмечают, что подобная интерпретация дает возможность суду определять в каждом конкретном случае, является ли та или иная система искусственного интеллекта источником повышенной опасности¹⁴. Заметим, что по умолчанию они таковыми не являются, поскольку большинство из них (по крайней мере тех, что используются чаще всего в повседневной жизни, например, голосовых помощников) не несет подобного рода рисков и не используется в рамках риск-ориентированного подхода к правовому регулированию. При этом среди недостатков такого подхода следует отметить факт наличия у большинства систем искусственного интеллекта автономности и способности к самообучению, что ставит безопасность их использования в прямую зависимость от действий разработчика, нежели пользователя (владельца) системы¹⁵. Необходимо также указать, что если искусственный интеллект не находится под контролем со стороны пользователя или владельца, то безусловное возложение на него ответственности не достигает одной из своих главных целей – побудить пользователей к поведению, обеспечивающему безопасность соответствующих систем¹⁶.

Концепция источника повышенной опасности не может отразить всю многогранность общественных отношений, которые складываются вокруг использования искусственного интеллекта, включающую в себя широкий круг субъектов, их различный вклад в его функционирование, определенный уровень автономности и возможность к самообучению, сложность технической конструкции и т.д.

В случаях, когда в действиях кого-либо (например, пользователь, генерирующий противоправный визуальный контент без его дальнейшего распространения) отсутствует состав правонарушения и отсутствовала возможность повлиять на алгоритм работы искусственного интеллекта в противоправных целях, вопрос ответственности должен рассматриваться индивидуально.

¹³ См.: постановление Пленума Верховного Суда РФ от 26.01.2010 № 1 «О применении судами гражданского законодательства, регулирующего отношения по обязательствам вследствие причинения вреда жизни или здоровью гражданина» // Официальный интернет-портал правовой информации (pravo.gov.ru) (дата обращения: 14.07.2024).

¹⁴ См.: Ижаев О.А., Кутейников Д.Л. Системы искусственного интеллекта и внедоговорная гражданско-правовая ответственность: риск-ориентированный подход // Lex russica. 2024. Т. 77. № 6. С. 23–34.

¹⁵ См.: там же.

¹⁶ См.: Buiten M., Strel A., Peitz M. The law and economics of AI liability // Computer Law & Security Review. 2023. Vol. 48. P. 9.

При этом стоит учитывать степень потенциальной повышенной опасности использования той или иной системы искусственного интеллекта вне зависимости от того, повлекло ли данное использование общественно опасные последствия. Суды при определении степени опасности использования системы искусственного интеллекта (и сообразно ответственности) также должны учитывать возможность владельца или пользователя системы контролировать и оперативно устранять потенциальные пробелы в ее деятельности, способные настичь общественно опасный вред на основании специфики той или иной системы ИИ, а также ее функционала.

При этом в сфере применения искусственного интеллекта уже выделяются наиболее критические и разрушительные риски в различных сферах – например, представленные в исследовании Центра безопасности в сфере ИИ (Cornell University)¹⁷. Карта таких рисков, представляется, может быть использована и для формирования правовых конструкций составов правонарушений с использованием ИИ-систем, а также конкретизации юридической ответственности за их противоправное с точки зрения национального и международного права применение.

В частности, выделяются следующие «серые зоны» в разработке и использовании систем искусственного интеллекта:

1. «ИИ-гонка» («гонка вооружений» среди государств и корпораций в целях технологического развития и автоматизации процессов), способствующая развитию конкуренции в сфере ИИ-технологий, а значит, и их совершенствованию. При этом возникает риск значительного увеличения их влияния на общество при стремлении государств и корпораций максимизировать выгоду в конкурентной борьбе.

К способам уменьшения рисков относительно данного сценария авторы относят: проактивное нормативное регулирование в сфере ИИ-безопасности, кибербезопасность, прозрачность и подотчетность данных и алгоритмов, контроль со стороны человека за решениями, принимаемыми подобными системами, в том числе общественный контроль над их функционированием, международное сотрудничество.

В то же время следует отметить, что возникает целый ряд вопросов, которые требует соответствующего решения: какие стандарты кибербезопасности будут применяться для предотвращения негативного воздействия такой конкуренции на общественную жизнь? Как будет обеспечиваться

¹⁷ См.: Hendrycks D., Mazeika M., Woodside T. An Overview of Catastrophic AI Risks [Электронный ресурс]. URL: <https://arxiv.org/pdf/2306.12001.pdf> (дата обращения: 15.07.2024).

подотчетность данных, когда речь идет о гонке технологий между двумя суверенными государствами, конкурирующими на ИИ-рынке? Кто и в каком порядке в этом случае будет осуществлять общественный контроль? В том числе остается не решенным вопрос и в части юридической ответственности: какой вид юридической ответственности применим к государствам или корпорациям, в действиях которых в условиях развития систем искусственного интеллекта усматриваются признаки тех или иных правонарушений?

2. Угрозы биобезопасности, среди возможных путей уменьшения которых авторы выделяют удаление из базовых моделей всей релевантной биоинформации, а также контроль за их утечками, ограничение доступа к наиболее продвинутым моделям и их инфраструктуре со стороны пользователя, выявление аномалий в функционировании систем ИИ, установление юридической ответственности для разработчиков систем ИИ общего назначения.

3. Организационные риски, которые могут приводить к разного рода катастрофическим событиям (техногенные катастрофы, крушения космических кораблей, самолетов, судов). В сфере использования искусственного интеллекта опасность возможной ошибки куда больше, нежели в поле деятельности человека, ввиду недостаточной проработанности правил поведения и протоколов безопасности систем ИИ в данной сфере деятельности.

Уменьшение рисков возникновения данного рода негативных последствий авторы предлагают за счет установления комплексных мер безопасности и, в том числе, введения обязательных экспертных проверок, внешнего независимого аудита и систем управления рисками. Следует отметить, что проблема прозрачности функционирования искусственного интеллекта является одной из наиболее сложных в части его нормирования и применения¹⁸. В этом случае возникает вопрос об ответственности аудиторов и экспертов, действия/бездействие которых могут допустить возникновение подобных рисков, а также о том, кем и как будут оцениваться такие системы управления рисками.

4. «Бунт» со стороны систем ИИ. Это может произойти через постепенное размывание и смещение заложенных человеком целей их функционирования или же в результате резкой потери контроля или обмана со стороны систем ИИ.

Данные риски можно уменьшить, например, за счет ограничения наиболее высокорисковых сценариев их использования, существования

¹⁸ См., напр.: Кутейников Д.Л., Ижаев О.А., Зенин С.С., Лебедев В.А. Алгоритмическая прозрачность и подотчетность: правовые подходы к разрешению проблемы «черного ящика» // Lex russica. 2020. Т. 73. № 6 (163). С. 139–148.

«красной кнопки», а также открытого допуска исследователей к базовым моделям.

При этом возникает новый риск – риск недобросовестного или противоправного допуска к таким моделям и их попадание в руки физических или юридических лиц, способных использовать данные модели и их модификации для совершения правонарушений. Наличие «красной кнопки» также подразумевает ответственность за ее противоправное использование, например, с целью сознательной остановки функционирования системы и ее развития в своих интересах с тем, чтобы выиграть время либо ресурсы для недобросовестной конкуренции в рамках «ИИ-гонки». Актуальным также остается вопрос о необходимости разработки протоколов и порядка применения «права красной кнопки» с закреплением этих моментов в национальном и международном законодательстве.

5. Вредное использование, заключающееся в возможности целенаправленно злонамеренно использовать системы ИИ. Среди таких выделяется биотerrorизм (создание новых вирусов, химического и биологического оружия), создание опасных ИИ-агентов, использование системы ИИ для манипуляции людьми, дезинформации и захвата власти. Данный список можно продолжить, поскольку категория «злонамеренного использования» может быть применена к работе практически любой системы искусственного интеллекта и будет отличаться лишь масштабом его последствий и уровнем их общественной опасности, а значит, и уровнем ответственности лица или группы лиц.

Отметим также, что указанные риски, последствия и пути их преодоления носят наднациональный характер, поскольку корпорации (а тем более государства) взаимодействуют на рынке ИИ-технологий на международном уровне, что порождает как конкуренцию, так и возможность сотрудничества. Но так или иначе, все эти процессы касаются не только внутринационального, но и наднационального, международно-правового регулирования, а значит, вопрос ответственности необходимо рассматривать также и на этом уровне.

При этом самым спорным и дискуссионным на данный момент является *вопрос о применении юридической ответственности к самому искусственному интеллекту как к субъекту права*. Здесь вопрос однозначно выходит за пределы правового исследования, затрагивая философские, социологические, психологические и иные научные проблемы, рассматриваемые в рамках общественных наук.

С учетом обозначенных выше подходов отметим, что в контексте исследования возможности применения юридической ответственности к системам искусственного интеллекта целесообразно говорить о таких важных для понимания ответственности

аспектов, как мотивация и целеполагание. Оба аспекта свойственны человеку, при этом не являются очевидными для систем искусственного интеллекта, действия которого изначально алгоритмизированы и определены человеком.

В то же время целеполагание в функционировании подобных систем нельзя отрицать, поскольку любой закладываемый в него алгоритм имеет конечную оформленную цель. Что касается мотивации, то ее спорность у искусственного интеллекта предопределена отсутствием биологических и социальных потребностей, обуславливающих мотивационный механизм на уровне психики. При этом здесь возникает «проблема согласования», о которой речь пойдет ниже.

Кроме того, говоря о предпосылках юридической ответственности в отношении систем искусственного интеллекта, следует остановиться на вопросе об основаниях юридической ответственности. Как было отмечено ранее, среди такого рода оснований очевидно к системам ИИ применимы общественная опасность и противоправность. При этом оставшиеся – виновность, интеллектуально-волевой характер и наказуемость – вызывают определенного рода сомнения.

Правонарушение – результат деятельности деликтоспособного субъекта, способного направлять и контролировать свою волю и поведение, рефлексируя и отдавая себе отчет в своих действиях, будучи способным нести ответственность за их пагубные последствия.

В условиях развития и жизнедеятельности «слабого» ИИ, выполняющего лишь узкий круг задач под руководством человека согласно заданным алгоритмам (голосовые и виртуальные помощники, автоматизированные системы, чат-боты, роботизированная техника, системы «умный дом» и др.), нельзя говорить об ответственности ИИ как деликтоспособного субъекта, поскольку в его действиях отсутствуют такие компоненты, как свобода воли, эмоциональный интеллект, понимание противоправности своих действий.

Наказуемость «слабого» ИИ как субъекта также не представляется возможной, поскольку он неспособен осознавать и принимать негативные последствия от применяемых санкций, не имея, к тому же, для этого устойчивой ценностной почвы, на которую эти санкции могут потенциально воздействовать. Соответственно, рассматривая категорию юридической ответственности в отношении искусственного интеллекта, можно говорить только об ответственности «сильного» искусственного интеллекта (*artificial general intelligence*, *artificial super intelligence*), который помимо общих признаков также обладает возможностью автономного принятия решений и способностью

к самосовершенствованию и самовоспроизведению. Данный процесс может быть оценен в качестве некой правовой фикции свободы воли, имеющей внешние формальные признаки, но отличные от человеческой воли (детерминированное психическими и нейронными процессами) происхождение и природу.

Поскольку на данный момент к ИИ невозмож но применить традиционную для отечественного и зарубежного права модель юридической ответственности, мы видим необходимым рассматривать данные системы в качестве принципиально иного субъекта, для которого категория «ответственность» и механизм ее реализации будут носить отличный с точки зрения своей сущности и назначения характер.

Здесь следует отметить, что ряд исследователей относят ИИ к т.н. квазисубъектам права (наравне с семьей, нацией, трудовым коллективом). Так, Е.В. Пономарева полагает, что роботизированные автономные системы не способны быть субъектами права, так как они «не обладают волей в правовом смысле, не имеют собственных интересов, не могут формировать субъективное отношение к совершающимся действиям, нести юридическую ответственность по причине отсутствия чувства вины»¹⁹. При этом автор указывает, что «акт признания законодателем их правоспособности не сделает их настоящими субъектами права, но может причинить вред праву как коммуникативной системе», поскольку такой квазисубъект так или иначе является стороной общественных отношений как источник возникновения юридического факта, в частности правонарушения.

Среди перечисленных автором подходов к пониманию субъекта прав (а значит, и юридической ответственности как части правового статуса такого субъекта) наиболее актуальным относительно исследования ответственности ИИ нам представляется постмодернистский подход, основанный на трудах Р. Барта, П. Бурдье и М. Фуко, получивший продолжение в отечественных работах В.И. Павлова, А.В. Полякова, И.Л. Честнова.

С их точки зрения, субъект права, с одной стороны, – это центр социальной реальности и объективного права, наделяющего его право- и дееспособностью. Он есть часть социального конструкта, вследствие чего является «неустойчивым, нецелостным и неразумным». С другой стороны, современный субъект права сам является создателем правовой реальности, будучи способным создавать и изменять ее своими действиями. Недостаток такого подхода, по мнению

¹⁹ Пономарева Е. В. Субъекты и квазисубъекты права: теоретико-правовые проблемы разграничения: дис. ... канд. юрид. наук. Екатеринбург, 2019. С. 11.

Е. В. Пономаревой, — отождествление субъекта права исключительно с человеком и, как следствие, отсутствие в праве рациональности²⁰. Причину возникновения категории «квазисубъект права» автор видит в размытости представлений о субъекте права в общественном сознании, их недостаточной детализации, а также доминировании в правовой науке формально-юридического подхода, не учитывающего сущностные и содержательные признаки данного явления.

Применение к системам искусственного интеллекта категории юридической ответственности усложняется отсутствием у последнего такого критерия определения субъекта права, как «правовая индивидуализация», представляющая собой требование к лицу для его распознаваемости, препятствующие отождествлению разных субъектов права, а именно определение имени, возраста для физических лиц, фирменного наименования, организационно-правовой формы, регистрации товарных знаков, места нахождения для юридических лиц²¹.

Краеугольным камнем современных исследований ИИ в сфере юридической ответственности является «проблема согласования». Ник Бостром описывает ее как способность «искусственного суперинтеллекта», будучи автономным и саморазвивающимся, достигать формально поставленных человеком целей при помощи нестандартных и изначально не предусмотренных человеком решений, способных нести общественно вредный противоправный характер (т.н. проблема извращенной реализации)²².

Следствием решения данной проблемы должно быть соответствие представлений о правовых нормах со стороны систем искусственного интеллекта аналогичным представлениям, принятым в обществе и государстве. В то же время на современном этапе это трудно реализуемо ввиду гибкости и приспособляемости общественного и индивидуального человеческого сознания к изменяющимся внешним условиям с пониманием контекста происходящих событий и отсутствия такового у искусственного интеллекта.

При этом ряд исследователей прогнозируют возможность избежать эффекта «извращенной реализации» с помощью точной алгоритмизации ИИ и максимально детализированной постановки задачи, однако на практике это не представляется возможным ввиду: а) непонимания глубоких внутренних алгоритмов работы ИИ и ее принципов в условиях самообучения; б) неспособности согласовать понимание потребностей человека с потребностями ИИ ввиду отсутствия таковых у последнего; в) невозможности

централизованно и согласованно контролировать данные процессы, особенно на международном уровне.

В совместном отчете «Злонамеренное использование ИИ: прогнозирование, предотвращение и смягчение последствий» (2018) был проанализирован спектр потенциальных угроз безопасности от злонамеренного использования технологий ИИ, где прямо было заявлено, что, если в ближайшее время не будут разработаны адекватные средства защиты, ИИ будет угрожать нашей цифровой, физической и политической безопасности. Американские специалисты Венделл Уоллах и Колин Аллен предположили, что по мере того, как роботы будут брать на себя все большую ответственность, они должны быть запрограммированы на способность принимать моральные решения ради нашей собственной безопасности²³. Авторы отчета говорят именно о перспективной юридической ответственности, требующей осознания и волевого анализа собственного потенциального правомерного поведения, при этом умалчивая о том, каким образом будет формироваться «моральность» таких решений у роботов, чем и как она будет обеспечена и каким образом будет караться нарушение подобной установки.

Оппоненты же данного подхода исключают возможность наделения ИИ и его носителей правосубъектностью. Согласно п. 3.2 Кодекса этики в сфере искусственного интеллекта, принятого в России Альянсом в сфере искусственного интеллекта и носящего рекомендательный характер, «акторы ИИ не должны допускать передачи полномочий ответственного нравственного выбора СИИ (систем искусственного интеллекта. — Прим. авт.), делегировать ответственность за последствия принятия решений СИИ — за все последствия работы СИИ всегда должен отвечать человек (физическое или юридическое лицо, признаваемое субъектом ответственности в соответствии с действующим законодательством Российской Федерации). Акторм ИИ рекомендуется принимать все меры для определения ответственности конкретных участников жизненного цикла СИИ с учетом их роли и специфики каждого этапа»²⁴.

Представляется, что сама по себе возможность применения юридической ответственности к системам искусственного интеллекта, тем более в условиях стремительно развивающихся технологий, связана с отступлением от устоявшегося концептуального подхода в классической юридической науке,

²⁰ См.: Пономарева Е.В. Указ. соч. С. 33.

²¹ См.: там же. С. 70.

²² См.: Bostrom N. Superintelligence: Paths, Dangers, Strategies. First. Oxford, 2014.

²³ См.: Гранин Р.С. Реф. ст.: Сутроп М. Проблемы согласования искусственного интеллекта с человеческими ценностями // Социальные и гуманитарные науки. Отечественная и зарубежная литература. Сер. 3. Философия. Реферативный журнал. 2021. № 3.

²⁴ Кодекс этики в сфере ИИ [Электронный ресурс]. URL: <https://ethics.a-ai.ru> (дата обращения: 14.07.2024).

который предполагает применение уже устоявшихся категориально-критериальных матриц к новым явлением объективной реальности. В этом смысле нарастающая необходимость решения вопроса о применении к системам искусственного интеллекта юридической ответственности предполагает необходимость выработки и нормативно-правового закрепления формально-определенных критериев такой ответственности.

В этой связи для наделения искусственного интеллекта возможностью быть субъектом юридической ответственности необходимо выполнение следующих условий²⁵:

1. Наличие алгоритма, наделяющего системы искусственного интеллекта возможностью осознавать причиняемый им вред и его характер. Важно понимать ключевые характеристики такого алгоритма: а) такое осознание должно коррелировать с моральными, этическими и правовыми нормами, действующими в обществе; б) понимание противоречия своих действий должно быть автономным, не требующим постоянного контроля со стороны человека и (или) какого-либо вмешательства или корректировок с его стороны; в) алгоритм должен давать понимание контекста ситуации с тем, чтобы обеспечить гибкость его действий в конкретных ситуациях: ИИ должен понимать, когда причинение вреда человеку является нарушением прав, а когда — крайней необходимостью, с полным пониманием ее границ, вне пределов которых наступает юридическая ответственность. При этом самосовершенствование без участия человека (как характеристика сильного искусственного интеллекта) также должно соответствовать развитию общественных норм (включая правовые) и принципам их понимания и соблюдения.

2. Наличие алгоритма, наделяющего системы искусственного интеллекта осознаваемой системой индивидуальных ценностей, доступных к стремлению их оберегать согласованными способами. В этом случае угроза таким ценностям (коими для человека являются жизнь, свобода, собственность, индивидуальные достижения и др.) дает возможность воздействия на искусственный интеллект со стороны человека и (или) правоприменительных органов через обратную связь. Однако данный вопрос упирается в невозможность применения человеческих ценностей, обусловленных биологической и социальной природой человека, к юридическому статусу ИИ, имеющего цифровую природу, к тому же до конца не изученную.

²⁵ Такого рода критерии применяются к системам сильного искусственного интеллекта, которые рассматриваются в качестве потенциальных субъектов юридической ответственности.

Однако следует отметить, что в настоящее время все алгоритмы создаются человеком (даже если код «пишет» ИИ, он делает это на основании заложенных человеком исходных данных). Соответственно, алгоритмы, наделяющие системы искусственного интеллекта возможностью осознавать причиняемый им вред и его характер и наделяющие системы искусственного интеллекта осознаваемой системой индивидуальных ценностей, доступных к стремлению их оберегать согласованными способами, также будут кем-то изначально созданы, и на данном этапе развития технологий не снимает суждения о том, что за действиями ИИ всегда стоит человек.

Несмотря на то что появление ИИ подобного типа учеными прогнозируется лишь к 2040—2050 гг.²⁶, необходимость выработки концепции регуляторного воздействия на него есть уже сейчас, поскольку развитие права и институтов его применения будет заметно отставать от развития ИИ, алгоритмы и принципы которого даже сейчас до конца не изучены, а представления о них не систематизированы. Самым действенным способом контролировать развитие ИИ (помимо протоколов безопасности, утверждаемых его разработчиками) на данный момент с точки зрения права является выработка принципиально новой юридической модели наделения его способностью нести ответственность и претерпевать негативные санкции со стороны человека.

Следует отметить, что внедрение подобных ценностей для его нормативного контроля является, по сути, юридическим конструктом, поскольку понимание ценностей человека и общества неприменимо к искусственно созданным системам, которым природа данных ценностей (свобода, социальное одобрение, материальные ценности) несвойственна.

3. Определение субъекта, управомоченного на применение юридической санкции. Говоря о традиционном понимании этой категории, нет сомнений в том, что санкцию к правонарушителю может применить только уполномоченный государственный орган. Такой порядок обусловлен тем, что государство обладает исключительной монополией на применение принудительных мер ввиду того, что оно представляет собой устойчивый институциональный механизм, обладающий равнодействующими инструментами и механизмами принуждения и осуществляющий его в установленных законом рамках.

Это направлено на недопущение общественного произвола и социального хаоса, вызванного неоднозначностью понимания социальных ценностей разными людьми, каждый из которых в своем желании установить или восстановить социальную справедливость может использовать совершенно разные, как аморальные, так и откровенно противоправные,

²⁶ См.: Bostrom N. Op. cit.

способы. Таким образом реализуется принцип равенства граждан перед законом и правосудием, а спектр санкций при этом также исчерпывающе установлен законодательством. Данное положение уравнивает граждан в парадигме «человек – государство», не допуская общественный произвол.

В то же время в отношении систем искусственного интеллекта допустимо говорить о возможности быть подвергнутыми санкциями не только со стороны государства, но и со стороны других субъектов (физических и юридических лиц: компании разработчика в лице ее руководящего органа, уполномоченного специалиста по машинному обучению и др.). Квазисубъектная сущность искусственного интеллекта вызывает необходимость создания специального органа, уполномоченного применять санкцию, тем самым оказывая сдерживающее воздействие на его деятельность. И если в системе взаимодействия «человек – человек» применение такого института невозможно (в противном случае государство потеряет свою монополию на право-применение) ввиду равенства физической природы и природы потребностей сторон, то в случае с ИИ такое представляется допустимым и даже разумным, поскольку даст специалистам в области ИИ и руководителям компаний определенные полномочия, сняв необходимость их реализации с органов государства и судебной системы.

Так или иначе, модель взаимодействия «человек – ИИ» в контексте наделения последнего правосубъектностью и полноценным правовым статусом будет так или иначе носить квазисубъектный характер в силу того, что юридическая ответственность человека связана с воздействием государства на его социальную и природную составляющую, в то время как для ИИ такая составляющая (даже при ее успешном внедрении в технологии работы ИИ) будет носить искусственный характер и будет ограничена технологическими возможностями в конкретный момент исторического развития технологий, науки и общества.

Юридической науке и практике предстоит решить множество юридических коллизий и нюансов, касающихся как правосубъектности систем искусственного интеллекта, так и возможной их деликтоспособности. Обозначенные критерии нуждаются в дальнейшей детализации с учетом поступательного развития общественных отношений в сфере применения ИИ, а также с учетом специфики различной отраслевой принадлежности. Тем не менее подобные критерии могут использоваться в качестве юридических оснований, позволяющих так или иначе говорить о применении юридической ответственности к профессиональному интеллекту.

СПИСОК ЛИТЕРАТУРЫ

1. Витрук Н.В. Общая теория юридической ответственности. М., 2008.
2. Гранин Р.С. Реф. ст.: Сутроп М. Проблемы согласования искусственного интеллекта с человеческими ценностями // Социальные и гуманитарные науки. Отечественная и зарубежная литература. Сер. 3. Философия. Реферативный журнал. 2021. № 3.
3. Ижаев О.А., Кутейников Д.Л. Системы искусственного интеллекта и внедоговорная гражданско-правовая ответственность: риск-ориентированный подход // Lex russica. 2024. Т. 77. № 6. С. 23–34.
4. Йонас Г. Принцип ответственности. М., 2004.
5. Кутейников Д.Л., Ижаев О.А., Зенин С.С., Лебедев В.А. Алгоритмическая прозрачность и подотчетность: правовые подходы к разрешению проблемы «черного ящика» // Lex russica. 2020. Т. 73. № 6 (163). С. 139–148.
6. Лейст О.Э. Санкции в советском праве. М., 1962. С. 47, 48.
7. Матузов Н.И., Малько А.В. Теория государства и права. М., 2001. С. 432.
8. Нерсесянц В. С. Общая теория права и государства. М., 1999.
9. Новиков Д.А. Трудовое право и искусственный интеллект: точки соприкосновения и расхождения // Ежегодник трудового права. 2024. № 14. С. 156–174.
10. Пономарева Е. В. Субъекты и квазисубъекты права: теоретико-правовые проблемы разграничения: дис. ... канд. юрид. наук. Екатеринбург, 2019. С. 11, 33, 70.
11. Савельев Ю.М. О понятии юридической ответственности // Юрид. исследования. 2015. № 10. С. 61–80.
12. Сырых В.М. Теория государства и права: учеб. М., 2012.
13. Харт Г. Понятие права. СПб., 2003.
14. Харт Г.Л.А. Приписывание ответственности и прав / пер. с англ. В.В. Оглезнева // Правоведение. 2010. № 5 (292). С. 116–135.
15. Чистое учение о праве Ганса Кельзена // К XIII конгрессу Международной ассоциации правовой и социальной философии (Токио, 1987): сб. переводов / пер. С.В. Лезова; отв. ред. В.Н. Кудрявцев, Н.Н. Разумович. М., 1987. Вып. 1. С. 164–166.
16. Bostrom N. Superintelligence: Paths, Dangers, Strategies. First. Oxford, 2014.
17. Buiten M., Strel A., Peitz M. The law and economics of AI liability // Computer Law & Security Review. 2023. Vol. 48. P. 9.
18. Hart H. L. A. The Ascription of Responsibility and Rights // Proceedings of the Aristotelian Society. 1949. Vol. 49. Pp. 171–194.
19. Hendrycks D., Mazeika M., Woodside T. An Overview of Catastrophic AI Risks [Электронный ресурс]. URL: <https://arxiv.org/pdf/2306.12001> (дата обращения: 15.07.2024).
20. Ingarden R. Kniężeczka o człowieku. Kraków, 1972.
21. Ingarden R. Über die Verantwortung. Ihre ontischen Fundamente. Stuttgart, 1970.

REFERENCES

1. *Vitruk N.V.* The general theory of legal responsibility. M., 2008 (in Russ.).
2. *Granin R.S.* Ref. art.: Sutrop M. Problems of matching artificial intelligence with human values // Social and humanitarian sciences. Domestic and foreign literature. Series 3. Philosophy. An abstract journal. 2021. No. 3 (in Russ.).
3. *Izhaev O.A., Kuteynikov D.L.* Artificial intelligence systems and non-contractual civil liability: a risk-oriented approach // Lex russica. 2024. Vol. 77. No. 6. Pp. 23–34 (in Russ.).
4. *Jonas H.* The principle of responsibility. M., 2004 (in Russ.).
5. *Kuteynikov D.L., Izhaev O.A., Zenin S.S., Lebedev V.A.* Algorithmic transparency and accountability: legal approaches to solving the problem of the “black box” // Lex russica. 2020. Vol. 73. No. 6 (163). Pp. 139–148 (in Russ.).
6. *Leist O.E.* Sanctions in the Soviet Union law. M., 1962. Pp. 47, 48 (in Russ.).
7. *Matuzov N.I., Mal'ko A.V.* Theory of state and law. M., 2001. P. 432 (in Russ.).
8. *Nersesyan V.S.* General theory of law and state. M., 1999 (in Russ.).
9. *Novikov D.A.* Labor Law and artificial intelligence: points of contact and discrepancies // Yearbook of Labor Law. 2024. No. 14. Pp. 156–174 (in Russ.).
10. *Ponomareva E.V.* Subjects and quasi-subjects of law: theoretical and legal problems of differentiation: dis. ... PhD in Law. Yekaterinburg, 2019. Pp. 11, 33, 70 (in Russ.).
11. *Saveliev Yu. M.* On the concept of legal responsibility // Legal researches. 2015. No. 10. Pp. 61–80 (in Russ.).
12. *Syrykh V.M.* Theory of state and law: textbook. M., 2012 (in Russ.).
13. *Hart H.* The concept of law. SPb., 2003 (in Russ.).
14. *Hart H.L.A.* Attribution of responsibility and rights / transl. from English by V.V. Oglezneva // Law studies. 2010. No. 5 (292). Pp. 116–135 (in Russ.).
15. The pure doctrine of law by Hans Kelsen // To the XIII Congress of the International Association of Legal and Social Philosophy (Tokyo, 1987): collection of translations / transl. S. V. Lezova; res. ed. V. N. Kudryavtsev, N. N. Razumovich. M., 1987. Iss. 1. Pp. 164–166 (in Russ.).
16. *Bostrom N.* Superintelligence: Paths, Dangers, Strategies. First. Oxford, 2014.
17. *Buiten M., Strel A., Peitz M.* The law and economics of AI liability // Computer Law & Security Review. 2023. Vol. 48. P. 9.
18. *Hart H.L.A.* The Ascription of Responsibility and Rights // Proceedings of the Aristotelian Society. 1949. Vol. 49. Pp. 171–194.
19. *Hendrycks D., Mazeika M., Woodside T.* An Overview of Catastrophic AI Risks [Электронный ресурс]. URL: <https://arxiv.org/pdf/2306.12001.pdf> (accessed: 15.07.2024).
20. *Ingarden R.* Kniążeczka o człowieku. Kraków, 1972.
21. *Ingarden R.* Über die Verantwortung. Ihre ontischen Fundamente. Stuttgart, 1970.

Сведения об авторах

ЗЕНИН Сергей Сергеевич –
кандидат юридических наук, доцент,
профессор кафедры теоретических
и публично-правовых дисциплин,
директор Института государства и права
Тюменского государственного университета;
625000 г. Тюмень, ул. Ленина, д. 38
ORCID: 0000-0002-4520-757X

ЯПРЫНЦЕВ Иван Михайлович –
кандидат юридических наук,
доцент кафедры теоретических и
публично-правовых дисциплин
Института государства и права
Тюменского государственного университета;
625000 г. Тюмень, ул. Ленина, д. 38

ИВАНОВА Лилия Викторовна –
кандидат юридических наук,
доцент кафедры уголовно-правовых дисциплин
Института государства и права
Тюменского государственного университета;
625000 г. Тюмень, ул. Ленина, д. 38
ORCID: 0000-0001-5255-3182

Authors' information

ZENIN Sergey S. –
PhD in Law, Associate Professor,
Associate Professor
of Theoretical and Public Law Disciplines,
Head of Institute of State and Law
of Tyumen State University;
38 Lenina str., 625000 Tyumen, Russia

YAPRYNTSEV Ivan M. –
PhD in Law,
Associate Professor
of Theoretical and Public Law Disciplines,
Institute of State and Law of Tyumen State University;
38 Lenina str., 625000 Tyumen, Russia

IVANOVA Liliya V. –
PhD in Law,
Associate Professor of Criminal Law Disciplines
of Institute of State and Law
of Tyumen State University;
38 Lenina str., 625000 Tyumen, Russia