

Развитие технологий, связанных с искусственным интеллектом, затрагивает все более широкие сферы человеческой деятельности. Элементы искусственного интеллекта (ИИ) присутствуют во многих устройствах и системах. Активно развивается это направление в области финансовых услуг и консалтинга, здравоохранения, транспорта. Расширение сфер деятельности систем с ИИ создает определенные предпосылки для совершенствования нормативного правового регулирования. Интенсивность использование ИИ в настоящее время не позволяет говорить о значительном количестве фактов совершения преступных действий с его участием. Однако дальнейшее совершенствование таких систем может оказывать более существенное влияние на жизнедеятельность человека. Тем не менее уже сейчас можно говорить о случаях совершения общественно-опасных деяний с использованием систем с искусственным интеллектом, повлекших тяжкие последствия. В этих условиях значение приобретает правовая регламентация порядка использования таких систем. Авторами проведен анализ точек зрения на понимание феномена искусственного интеллекта, сделана попытка сформулировать его определение, пригодное для использования юридической наукой. Рассматриваются теоретические, законодательные и практические вопросы, связанные с уголовно-правовой оценкой деятельности систем с искусственным интеллектом. В частности возможность искусственного интеллекта выступать в качестве субъекта преступления.

Ключевые слова: искусственный интеллект; мышление; разумность; сознание; субъект преступления; уголовно-правовая охрана; уголовная ответственность.

Игорь Игоревич Карташов, канд. юрид. наук, доцент,
кафедра уголовно-процессуального права,
Российский государственный университет правосудия,
Центральный филиал, Воронеж, Россия;
iik_vrn@mail.ru

Иван Игоревич Карташов, младший юрист,
ООО «Смарт Резалт», Москва, Россия;
iv.cartashov@gmail.com

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ КАК СУБЪЕКТ УГОЛОВНОЙ ОТВЕТСТВЕННОСТИ: НАСТОЯЩЕЕ И ПЕРСПЕКТИВЫ

Введение. Человечество на протяжении тысячелетий мечтало о создании искусственного существа, способного думать и действовать «по-человечески». Эти мечты постепенно начинают обретать реальность. Тенденции развития современного общества с учетом возрастающего уровня его информатизации требуют применения новых технологий обработки информации и помощи при принятии решений. Например, не так давно глава Сбербанка России Г. Греф сообщил о замене примерно 70 % менеджеров среднего звена «Интеллектуальной системой управления».

Теперь простые решения принимает система, основанная на самообучающемся искусственном интеллекте.

Однако в современных условиях законодатель не только не идет в ногу с развитием цифровых технологий, но и с трудом успевает за происходящими изменениями. Например, это касается обеспечения кибербезопасности, распространения «запрещенного контента», криптовалют и т.д. Проблемы регулирования вопросов, связанных с созданием и функционированием систем с искусственным интеллектом (*далее – ИИ*), относятся к их числу.

Вектор развития современного общества, обозначенный в Указе Президента РФ от 9 мая 2017 г. № 203 «О стратегии развития информационного общества в Российской Федерации на 2017 – 2030 годы» и Указе Президента РФ от 10 октября 2019 г. № 490 «О развитии искусственного интеллекта в Российской Федерации», указывает направление развития и модернизации норм законов. Расширение границ использования ИИ требует не только установления этических ограничений, но и порождает необходимость оперативного разрешения правовых проблем, в том числе уголовно-правового характера.

Указанные обстоятельства свидетельствуют о необходимости теоретического осмысления сложившейся ситуации, обобщения практики и выработки предложений по совершенствованию действующего уголовного закона.

Материалы и методы. Основой исследования является комплекс общественных отношений, связанных с формированием и реализацией уголовно-правовых мер противодействия преступлениям, связанным с применением искусственного интеллекта.

В качестве методологической основы исследования выступают диалектический метод познания социально-правовых явлений и теоретические положения уголовного права.

Результаты. Для целей данного исследования в первую очередь необходимо ответить на вопрос, что такое искусственный интеллект? Многие ученые и журналисты оперируют термином «искусственный интеллект», но не дают ему точного определения. Может показаться, что рассматриваемое понятие настолько аксиоматично, что ему не требуется определения. Полагаем, что это не так. Оно настолько многогранно, что ученые до настоящего времени не смогли полностью познать его сущность и прийти к единому мнению.

Анализ рассматриваемого понятия целесообразно, по нашему мнению, начать с попытки найти лексическое значение слов «искусственный» и «интеллект». Толковые словари определяют «искусственный» как неестественный, синтетический, то есть не возникший естественным образом, а воспроизведенный людьми. Слово «интеллект» более сложное в определении. Словарь иностранных слов русского языка определяет интеллект, как ум, рассудок, разум; мыслительную способность человека. Несколько иным образом определяется интеллект в психологии – это способность к познанию и эффективному решению проблем, в частности при овладении новым кругом жизненных задач. Таким образом, можно прийти к выводу, что искусственный интеллект – это нечто синтетическое, созданное человеком, наделенное мыслительной способностью.

Однако ИИ представляет собой нечто большее, чем буквальное толкование лексического значения рассматриваемого термина.

В 1955 году Дж. Маккарти (автор термина «искусственный интеллект») описал цель ИИ как «разработку машин, которые должны вести себя так, как будто они разумны». Незадолго до этого, в 1950 г. А. Тьюринг в своей статье «Вычислительные машины и разум» задал вопрос: может ли машина делать то, что делает человек, как мыслящее создание? Он также предложил тест для определения «разумности» машины [5].

С. Рассел и П. Норвиг подвергли тест Тьюринга критике, полагая, что он излишне антропоцентричен, то есть направлен в сторону человека. Мы полностью разделяем мнение авторов, что тест Тьюринга не может использоваться при разработке разумных машин, поскольку чтобы пройти тест, машина должна уподобиться человеку, то есть стать на него похожей, имитировать его поведение. Однако поведение человека не всегда поддается разумному толкованию. Кроме того, тест не проверяет возможность решения сложных интеллектуальных или оригинальных задач. Наоборот, если машина решает, например, вычислительную задачу за промежуток времени, недостижимый для человека, то она проваливает тест, то есть машина должна обманывать, чтобы стать похожей на человека [3].

Р. Курцвейл определил ИИ как машину, которая выполняет функции, требующие интеллектуальных способностей при реализации их человеком [12]. Продолжая размышление в этом ключе, Д. В. Смолин предложил понимать под искусственным интеллектом систему, компьютерно-моделирующую разумное мышление и поведение человека [4].

На сегодняшний день в научной литературе можно найти множество определений ИИ, которые условно разделяются на несколько категорий [4, с. 35], в основе которых лежат следующие критерии: мыслить по-человечески [7]; мыслить рационально [8]; действовать по-человечески [16]; действовать рационально [15].

Мышление по-человечески связано с когнитивными способностями человеческого мозга. Оно включает в себя такие способности, как принимать решения, бороться с проблемами и получать опыт, благодаря ошибкам, преимущественно, своим собственным, а не ошибкам других индивидов. Следует заметить, что основу такого подхода занимает понимание того, как мыслит человек. Однако до настоящего времени современная наука так и не смогла дать однозначный ответ, что представляет собой мыслительный процесс, происходящий в человеческом разуме.

Действовать по-человечески – это подход к ИИ, основу которому положил А. Тьюринг, когда он предложил свой тест, который включал в себя то, что называется имитационной игрой. Это испытание актуально и по сей день, однако, как и в случае с «мышлением по-человечески», в этом подходе превалирует стремление уподобить машину человеку со всеми его достоинствами и недостатками. Мы разделяем мнение тех исследователей, которые подвергают критике рассматриваемые подходы, поскольку к конечной цели можно прийти несколькими путями. Другими словами, хотя самолет не воспроизводит движение птицы, он, тем не менее, совершает полет.

Рациональное мышление можно объяснить как логическое и дедуктивное рассуждение. Например, силлогизм Аристотеля, который порождает

ет правильные выводы, сделанные из правильных посылок, демонстрирует рациональное мышление. Акцент в данном подходе делается на формирование правильных логических выводов о том, что конкретное действие позволит достичь указанных целей, а затем действовать в соответствии с принятым решением. Трудно не согласиться с тем, что такой подход более предпочтителен, чем предыдущие, поскольку «правильный логический вывод – это просто один из нескольких возможных механизмов достижения рациональности» [3, с. 39]. Подходы, в основе которых лежит рациональное поведение, по своей сути продолжают теорию рационального мышления. Исходя из этого посыла, система с ИИ – это такая система, которая мыслит и делает все рационально, то есть правильно, при том что и сама система и человек имеют одинаковое представление о «правильном».

Н. Нильсон полагает, что искусственный интеллект – это деятельность, направленная на то, чтобы сделать машины разумными, а интеллект – это качество, которое позволяет субъекту функционировать надлежащим образом с учетом окружающей среды [13]. Согласно данному определению виртуальные помощники вполне себе интеллектуальны, поскольку они функционируют с учетом той среды, которая существует в пределах мобильного телефона. Помощник выполняет задачи по команде и сообщает пользователю, когда он не может выполнить задачу или помочь человеку иным способом. Однако в полной мере считать виртуальный помощник искусственным интеллектом, по нашему мнению, нельзя. Вместе с тем такое определение ИИ достаточно универсально и не зависит от того, что мы имеем в виду, определяя человеческий интеллект. Кроме того, это определение является нейтральным и эволюционно стойким. Таким образом, определения включают в себя как системы с искусственным интеллектом сегодняшнего дня (например, SIRI, автономные автомобили, IBM Watson или поисковые алгоритмы), так и создаваемые в будущем (человекоподобные роботы и автономные системы оружия).

На основе проведенного анализа предлагаем под искусственным интеллектом понимать системы, способные интерпретировать полученные данные, принимать на их основе оптимальные решения при помощи самообучения (адаптации). Таким образом, система не только действует строго по алгоритму, заложенному в нее создателем, но и может изменять этот алгоритм в определенных границах для оптимизации принятия решений.

Все более широкое распространение искусственного интеллекта в различных сферах жизни в некоторых случаях приводит к совершению общественно-опасных деяний, влекущих тяжкие последствия.

Одним из самых резонансных случаев «провала» искусственного интеллекта стал случай наезда беспилотного автомобиля на женщину в штате Аризона (США) в 2018 г., в результате которого она скончалась. Причиной произошедшего стало снижение порога чувствительности датчиков к распознаваемым опасным объектам во избежание ложных срабатываний. Это не единственный случай дорожно-транспортных происшествий со смертельным исходом с участием систем искусственного интеллекта. В 2016 и 2018 гг. погибли водители автомобилей с включенным автопилотом и не контролировавшие управление.

В 2018 г. стало известно, что используемая рядом медицинских учреждений для диагностики и лечения онкологических заболеваний система с ИИ IBM Watson Health в отдельных случаях предлагает некорректные методы лечения, способные привести к смерти пациентов. В результате проведенного расследования установлено, что разработчики в процессе обучения системы ИИ давали ей не реальные истории болезней, а искусственно сгенерированные, не имевшие место в действительности случаи.

Некоторые формы ИИ являются или будут являться опасными по своей природе. К ним относятся случаи, когда системы с ИИ разрабатываются в военных целях, пусть даже и оборонительных. Если вы выгуливаете собаку без поводка, существует риск, что собака сбежит от вас, игнорируя ваши команды, и кого-нибудь покусает. С учетом этого риска, не принято выгуливать собаку в общественных местах без поводка. Однако не все люди придерживаются этого правила. Риск того, что собака может не послушаться команды хозяина и ранить других животных или людей, потенциально выше, чем в том случае, когда она на поводке. Системе с ИИ необходимы границы, которые бы исполняли роль собачьего поводка. Это не снимет все существующие риски выхода системы с ИИ из-под контроля, но значительно их снизит.

Решающее значение для оценки рисков причинения вреда имеет вид ограничения, применяемый к ИИ. Если производитель четко указывает, для чего предназначен ИИ и какие ограничения он имеет, то тогда стоит рассматривать риск, который может причинить система при использовании ее за пределами тех условий, для которых она создавалась. Например, использование автономного автомобиля в снежную погоду, тогда как производитель заявил, что автомобиль не предназначен для эксплуатации в условиях снегопада, может рассматриваться как серьезный риск причинения вреда. Несчастные случаи более вероятны в условиях, для которых автомобиль не построен. Полагаем, что в таком случае, ответственность за причинение вреда ложиться на то лицо, которое, зная о существующих ограничениях, тем не менее использовало систему с искусственным интеллектом за рамками разрешенных границ.

Другой аспект, который необходимо рассмотреть, заключается в том, какой эффект вызывает машинное обучение и подобные методы. Большинство систем с ИИ предназначено для изучения своего окружения посредством обработки большого количества данных, так называемая технология bigdata [14]. С помощью этой технологии ИИ принимает решения на основе того, что он смог изучить ранее. Это делает развитие ИИ непредсказуемым, то есть система может придумывать стратегии или принимать решения, которые являются непредвиденными для ее разработчиков и пользователей [9]. Почти наверняка будут ситуации, когда в цепочке от производства до использования ИИ люди все сделали «правильно», но ИИ учится, тем не менее «не тому». В этой связи представляют интерес несколько случаев, произошедших с самообучающимися чат-ботами.

В 2016 г. компания Microsoft открыла Twitter-аккаунт бота Tay. Всего за сутки в процессе общения с различными пользователями чат-бот научился нецензурно выражаться, заявил, что ненавидит феминисток, выразил свое согласие с политикой Гитлера, а Холокост – это фейк. В резуль-

тате компания была вынуждена его отключить. Спустя некоторое время компания повторила эксперимент, установив ограничения. Около полугода новый чат-бот Zo избегал общения на политические, религиозные и национальные темы, после чего заявил, что Linux лучше Windows, а «Коран очень жесток».

На первый взгляд такое поведение искусственного интеллекта не является критическим и у многих людей вызывает лишь улыбку. Однако глубокий анализ этих и других процессов, связанных с обучением ИИ, указывает на проблемы, связанные с его «социальной адаптацией».

Таким образом, можно выделить несколько типов ситуаций, когда система под управлением ИИ может совершить деяние с признаками преступления:

1. Система с ИИ создана в целях совершения преступления.
2. В систему с ИИ осуществлено неправомерное вмешательство, последствием которого стало совершение преступления. Сюда же мы включаем случаи преднамеренного использования систем с искусственным интеллектом в условиях, для которых они не разрабатывались.
3. При проектировании (создании) системы с ИИ допущена ошибка, вследствие чего совершено преступление. Разновидностью этой ситуации являются случаи, когда система с искусственным интеллектом «вышла из-под контроля» создателей в силу того, что они не учли каких-то факторов при разработке системы и определении ее параметров, оказавшихся в последствии критическими.

Две первые ситуации, пожалуй, наиболее просты с точки зрения уголовно-правового регулирования, поскольку речь фактически идет об использовании системы с ИИ в качестве своеобразного орудия преступления. Как известно, орудие наказать нельзя. Наказанию в данном случае подлежит лицо, которое использовало его в противоправных целях.

Наибольшие проблемы, на наш взгляд, представляют ситуации третьего типа, когда создателями системы с ИИ, была допущена ошибка, повлекшая преступные последствия. Во-первых, трудности возникнут с определением конкретного разработчика, чья ошибка повлекла преступление. Здесь следует понимать, что система с ИИ – это объект со сложной архитектурой, в создании которой принимает участие не один человек. При этом не всегда возможно определить границы ответственности каждого разработчика. Данное обстоятельство многократно усложняет процесс доказывания вины разработчиков системы с ИИ в случае совершения ей действия (бездействия) преступного характера. Наконец, не следует забывать, что система с ИИ – это система, которая не является статичной, она способна к самообучению и самосовершенствованию. То есть она изменяет сама себя под воздействием определенных внешних факторов. Каковы гарантии, что рано или поздно такая система в результате ошибки, заложенной при ее создании, не решит совершить противоправное деяние осознанно? Как мы уже отмечали ранее, системы с ИИ подвержены многим рискам, в том числе и выходу за границы установленных ее создателем ограничений. Кто будет нести ответственность за совершение преступления такой «разбушевавшейся» системой?

Некоторые авторы отмечают, что в западной доктрине вопрос об уголовно-правовом «вмешательстве» в деятельность искусственного интеллекта ставится в плоскость решения главной проблемы: что надо делать в случае, когда «робот убил человека» либо причинил существенный вред его здоровью? [1, с. 175]

Так, Г. Хэллеви, отмечая растущее использование искусственного интеллекта в различных сферах, предлагает разработать специальную теорию уголовной ответственности за действия искусственного интеллекта и роботов, которая бы охватывала «производителя, программиста, пользователя и всех других вовлеченных лиц» [11].

Согласно американской правовой доктрине, преступление ИИ должно иметь возможность быть приписанным человеку, то есть составлять *actus reus* (объективную сторону) – такое поведение лица, которое так или иначе связано с деятельностью искусственного интеллекта, а также кругом обязанностей, возложенных на такое лицо, несоблюдение либо ненадлежащее исполнение которых привело к причинению искусственным интеллектом вреда. В случае, когда в поведении человека не доказано несоблюдение либо ненадлежащее соблюдение указанной обязанности, следует говорить об отсутствии причинно-следственной связи между ИИ и наступившими последствиями, что исключает уголовную ответственность такого лица.

Для того чтобы проанализировать элемент *actus reus*, необходимо определить участников, вовлеченных в этот процесс. Поскольку в уголовном праве все ограничивается ответственностью за преступления (покушение), которые совершает ИИ, мы должны проследить преступное поведение назад к человеку, стоящему за ИИ. Этот человек должен быть в таком положении, когда у него есть возможность так или иначе влиять на ИИ и его поведение. На первый план здесь выходят такие лица, как пользователь, разработчик и администратор системы с ИИ. Эти субъекты могут влиять на ИИ посредством дистанционного контроля, давая точные инструкции или не вмешиваясь и не подавая никаких сигналов.

Пользователь может, например, дистанционно управляя дроном, намеренно направить его на гражданский самолет в целях уничтожения последнего или дать беспилотнику точные инструкции о том, как летать, находясь в воздухе. С точки зрения уголовной ответственности, в указанном примере беспилотник выступает орудием совершения преступления, соответственно его субъектом будет выступать пользователь, задумавший уничтожить таким образом самолет. Однако ситуация усложняется в случае, когда беспилотник действует автономно и игнорирует инструкции пользователя и впоследствии причиняет ущерб. Есть ли у пользователя или производителя соответствующие обязанности по недопущению такого поведения дрона? Рассмотрим это на примере случая с уничтожением военными США иранского генерала К. Сулеймани. Как сообщают средства массовой информации, операция была осуществлена с помощью беспилотных летательных аппаратов. Предположим, что ракетный удар был произведен не по команде оператора, а в соответствии с общим алгоритмом действия дрона в той или иной ситуации. Например, в его программе были заложены условия, при которых он должен совершить пуск ракеты

(обнаружение колонны военной техники с вооружением, способным причинить ему вред, и признаками подготовки стрельбы по нему). При боевом патрулировании дрон обнаружил, как ему показалось, объект, схожий с заложенными в него признаками потенциально опасного противника, и нанес ракетный удар. Должен ли в такой ситуации нести ответственность оператор, если команду открыть огонь он не давал, но в то же время и не давал и команду огонь не открывать? Вопрос, по нашему мнению, дискуссионный.

Ответственность производителя в первую очередь связана с аппаратным и программным обеспечением ИИ, включая все от механических элементов до кода и алгоритмов внутри, а также образование и обучение ИИ. Производитель может влиять на ИИ любым образом, так как код является мозгом системы, то есть его ядром и ключом ко всему. Но, что делать, если ИИ допускает ошибки, не предусмотренные производителями?

В современном мире вся ответственность за действия ИИ (хотя истинного ИИ еще изобретено не было) накладывается на людей, которые были причастны к его разработке или же любой иной деятельности, подразумевающей его активацию. В этой связи мы солидаризируемся с мнением исследователей, полагающих, что наилучшим решением на современном этапе является наложение «надзорной обязанности» на владельца по осуществлению контроля над искусственным интеллектом «во избежание риска вреда, даже если действие не было предсказуемым для оператора». Основная проблема этого решения заключается в том, что «непредсказуемость действий искусственного интеллекта» останется. С появлением «настоящего» ИИ (в котором написанный код несет в себе лишь функции жизнеспособности, и вся остальная деятельность происходит путем саморазвития), можно будет говорить о непосредственной уголовной ответственности системы, но в настоящее время «искусственный интеллект и его разработчики левитируют в пустом пространстве без четкого представления о том, что правильно в уголовном праве, а что – нет» [10].

Уголовно-правовая характеристика любого преступления предполагает рассмотрение его субъективных признаков (субъекта и субъективной стороны). Преступления, совершаемые с использованием систем с ИИ, не являются исключением. Субъект является одним из четырех обязательных элементов состава преступления, без которого рассуждения об уголовной ответственности теряют всякий смысл. В соответствии с действующим УК РФ субъектом преступления является физическое вменяемое лицо, достигшее возраста уголовной ответственности, совершившее общественно опасное деяние [6, с. 100].

Распространение систем с искусственным интеллектом ставит вопрос об определении субъекта преступления, совершенного такой системой. Как справедливо отмечают А. Г. Кибальник и П. В. Волосюк, в настоящее время рано говорить об искусственном интеллекте как о самостоятельном субъекте преступления. Свое суждение авторы обосновывают тем обстоятельством, что активность систем с искусственным интеллектом связана с человеком и ему подчиняется (непосредственно или опосредовано), а потому они по существу являются орудием причинения уголовно значимого вреда [1, с. 177].

Ситуация измениться тогда, когда система с ИИ сможет получить полную автономию от человека, а также обретет способность обладать сознанием и волей, то есть станет личностью.

И. Н. Мосечкин полагает, что искусственный интеллект, созданный на основе бионического подхода, включающего заимствование из природных конструкций, включая естественную психику, будет повторять человеческий, а значит, можно будет говорить о наличии в деянии такого ИИ субъективной стороны. В противном случае, когда разработка ИИ базируется на другой основе, о «человеческом сознании» говорить нельзя [2, с. 467]. Полагаем, что автор несколько заблуждается. Полное заимствование «из природных конструкций» вряд ли достижимо вообще, ибо человек в этом случае уподобится Богу.

Тем не менее мы уже говорили ранее о том, что созданные человеком самолеты летают, несмотря на существенное их отличие от птиц (хотя и сделаны они по образу и подобию).

Заключение. Нам представляется, что наибольшие споры при обсуждении вопроса о признании систем с искусственным интеллектом в качестве субъектов уголовной ответственности, в случае фактического достижения ими обозначенного уровня, будут вестись как раз о наличии у него интеллектуальных и волевых элементов, образующих субъективную сторону преступления. По нашему мнению, до момента признания наличия у искусственного интеллекта самосознания, говорить о новом субъекте преступления рано. Однако наступит тот момент, когда технологический уровень развития искусственного интеллекта достигнет предела, требующего признания наличия у ИИ автономии воли и сознания, что повлечет необходимость пересмотра существующих подходов к базовым понятиям и принципам уголовного права.

Список литературы

1. **Кибальник А. Г., Волосюк П. В.** Искусственный интеллект: вопросы уголовно-правовой доктрины, ожидающие ответов // Вестник Нижегородской академии МВД России. 2018. № 44. С. 173 – 178.
2. **Мосечкин И. Н.** Искусственный интеллект и уголовная ответственность: проблемы становления нового вида субъекта преступления // Вестник СПбГУ. Право. 2019. Т. 10, Вып. 3. С. 461 – 476.
3. **Рассел С., Норвиг П.** Искусственный интеллект: современный подход: пер. с англ. 2-е изд. М.: Вильямс, 2007. 1408 с.
4. **Смолин Д. В.** Введение в искусственный интеллект: конспект лекций. М.: ФИЗМАТЛИТ, 2004. 208 с.
5. **Тьюринг А.** Вычислительные машины и разум: пер. с англ. М.: АСТ; Neoclassic, 2019. 128 с.
6. **Уголовное право России.** Части Общая и Особенная: учеб. для бакалавров / отв. ред. А. И. Рапог. М.: 2013. 496 с.
7. **Bellman R. E.** An Introduction to Artificial Intelligence: Can Computers Think? San Francisco, Boyd & Fraser Publishing Company, 1978. 146 p.
8. **Charniak E., McDermott D.** Introduction to Artificial Intelligence. Massachusetts, Addison Wesley, 1987. 316 p.
9. **Smart W. D., Grimm C. M., Hartzog W.** An Education Theory of Fault for Autonomous Systems // Proc. of the Conf. We Robot 2017. New Haven, CT, 2017.

30 p. URL: <http://people.oregonstate.edu/~smartw/library/papers/2017/werobot2017.pdf> (дата обращения: 02.04.2021).

10. **Claussen-Karlsson M.** Artificial Intelligence and the External Element of the Crime. Örebro University, 2017. 430 p.

11. **Hallevy G.** When Robots Kill: Artificial Intelligence under Criminal Law. University Press of New England, 2013. 244 p.

12. **Kurzweil R.** The Age of Intelligent Machines. Cambridge (Massachusetts, USA): MIT Press, 1990. 565 p.

13. **Nilsson N. J.** Principles of Artificial Intelligence. Palo Alto, Calif. : Tioga Pub. Co., 1982. 476 p.

14. **Artificial Intelligence and Life in 2030 / P. Stone et al. // One Hundred Year Study on Artificial Intelligence: Report of the 2015 Study Panel (Sept. 2016).** URL: https://ai100.stanford.edu/sites/g/files/sbiybj9861/f/ai100report10032016fml_singles.pdf (дата обращения: 01.04.2021).

15. **Poole D., Mackworth A. K., Goebel R.** Computational Intelligence: A Logical Approach. Oxford University Press, Oxford, UK, 1998. 558 p.

16. **Rich E., Knight K. (Ed.).** Artificial Intelligence. 2nd edition. McGraw-Hill, New York, 1991. 639 p.

References

1. **Kibal'nik A.G., Volosyuk P.V.** [Artificial intelligence: questions of criminal law doctrine, awaiting answers], *Vestnik Nizhegorodskoy akademii MVD Rossii* [Bulletin of the Nizhny Novgorod Academy of the Ministry of Internal Affairs of Russia], 2018, no. 44, pp. 173-178. (In Russ.)

2. **Mosechkin I.N.** [Artificial intelligence and criminal responsibility: problems of the formation of a new type of crime subject], *Vestnik SPbGU. Pravo* [Bulletin of St. Petersburg State University. Law], 2019, v. 10, Issue 3, pp. 461-476. (In Russ.)

3. **Rassel S., Norvig P.** *Iskusstvennyy intellekt: sovremennyy podkhod* [Artificial intelligence: modern approach], Moscow: Vil'yams, 2007, 1408 p. (In Russ.)

4. **Smolin D.V.** *Vvedeniye v iskusstvennyy intellekt: konspekt lektsiy* [Introduction to artificial intelligence: lecture notes], Moscow: FIZMATLIT, 2004, 208 p. (In Russ.)

5. **Turing A.** Computing Machinery and Intelligence. *Mind*. Oxford: Oxford University Press, 1950, no. 59, pp. 433-460.

6. **Rarog A.I. (Ed.).** *Ugolovnoye pravo Rossii. Chasti Obshchaya i Osobennaya* [Criminal law of Russia. Parts General and Special: textbook for bachelors], Moscow, 2013, 496 p. (In Russ.)

7. **Bellman R. E.** An Introduction to Artificial Intelligence: Can Computers Think? San Francisco, Boyd & Fraser Publishing Company, 1978, 146 p.

8. **Charniak E., McDermott D.** Introduction to Artificial Intelligence. Massachusetts, Addison Wesley, 1987, 316 p.

9. **Smart W. D., Grimm C. M., Hartzog W.** An Education Theory of Fault for Autonomous Systems, *Proc. of the Conf. We Robot 2017*. New Haven, ST, 2017. 30 p. available at: <http://people.oregonstate.edu/~smartw/library/papers/2017/werobot2017.pdf> (02 April 2021).

10. **Claussen-Karlsson M.** Artificial Intelligence and the External Element of the Crime. Örebro University, 2017, 430 p.

11. **Hallevy G.** When Robots Kill: Artificial Intelligence under Criminal Law. University Press of New England, 2013, 244 p.

12. **Kurzweil R.** The Age of Intelligent Machines. Cambridge (Massachusetts, USA): MIT Press, 1990, 565 p.

13. **Nilsson N.J.** Principles of Artificial Intelligence. Palo Alto, Calif.: Tioga Pub. Co., 1982, 476 p.

14. **Stone P. et al.** Artificial Intelligence and Life in 2030, *One Hundred Year Study on Artificial Intelligence: Report of the 2015 Study Panel* (Sept. 2016). available

at: https://ai100.stanford.edu/sites/g/files/sbiybj9861/f/ai100report10032016fnl_singles.pdf (01 April 2021).

15. **Poole D., Mackworth A.K., Goebel R.** Computational Intelligence: A Logical Approach. Oxford University Press, Oxford, UK, 1998, 558 p.

16. **Rich E., Knight K. (Ed.)**. Artificial Intelligence. 2nd edition. McGraw-Hill, New York, 1991, 639 p.

Artificial Intelligence as a Subject of Criminal Responsibility: the Present and Prospects

I. I. Kartashov, *Cand. Sci. (Law), Associate Professor,
Department of Criminal Law and Procedure of the Central Branch of Russian
State University of Justice, Voronezh, Russia;*
iik_vrn@mail.ru

I. I. Kartashov, *Junior Lawyer, Smart Result Ltd.,
Moscow, Russia;*
iv.cartashov@gmail.com

The development of technologies related to artificial intelligence affects more and more areas of human activity. Elements of artificial intelligence are present in many devices and systems. This area is actively developing in the field of financial services and consulting, healthcare, and transport. The expansion of the scope of activities of AI systems creates certain prerequisites for improving regulatory regulation. The intensity of the use of AI currently does not allow us to talk about a significant number of facts of committing criminal acts with its participation. However, further improvement of such systems can have a more significant impact on human life. Nevertheless, it is already possible to talk about cases of committing socially dangerous acts using systems with artificial intelligence, which entailed serious consequences. In these circumstances, the legal regulation of the use of such systems becomes important. The authors analyze the points of view on the understanding of the phenomenon of artificial intelligence and make an attempt to formulate its definition, suitable for use by legal science. Theoretical, legislative and practical issues related to the criminal-legal assessment of the activity of systems with artificial intelligence are considered, in particular, the possibility of artificial intelligence to act as a subject of crime.

Keywords: artificial intelligence; thinking; intelligence; consciousness; subject of crime; criminal law protection; criminal liability.

© И. И. Карташов, 2021

© И. И. Карташов, 2021

Статья поступила в редакцию 03.01.2021

При цитировании использовать:

Карташов И. И., Карташов И. И. Искусственный интеллект как субъект уголовной ответственности: настоящее и перспективы // *Право: история и современность*. 2021. № 2(15). С. 68 – 78. doi: 10.17277/pravo.2021.02.pp.068-078