

НЕОБХОДИМОСТЬ ПРАВОВОГО РЕГУЛИРОВАНИЯ «ЧЕРНОГО ЯЩИКА» ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Головин Кирилл Сергеевич, аспирант, Институт законодательства и сравнительного правоведения при Правительстве РФ, адвокат Адвокатской палаты Московской области, Москва, Россия, golovinks@gmail.com, <https://orcid.org/0009-0002-8844-488X>

Аннотация. В данной статье рассматриваются вопросы прозрачности и регулирования искусственного интеллекта (ИИ). В ней обсуждается растущая тенденция использования алгоритмов принятия решений в различных областях жизни и необходимость в прозрачных и понятных системах ИИ. Автор освещает концепцию «черного ящика», где алгоритмы и их процессы становятся скрытыми для пользователей и разработчиков, и подчеркивает важность объяснимости и доступности информации о работе ИИ. Статья также затрагивает вопросы, касающиеся права на объяснение и возможных стратегий обеспечения прозрачности и подотчетности при принятии алгоритмических решений, включая документирование параметров и сертификацию систем.

Ключевые слова: прозрачность искусственного интеллекта, черный ящик ИИ, регулирование искусственного интеллекта, алгоритмическое принятие решений, правовые аспекты ИИ, сертификация систем ИИ, оспаривание решений ИИ.

Для цитирования: Головин К.С. Необходимость правового регулирования «черного ящика» искусственного интеллекта // Вестник Костромского государственного университета. 2024. Т. 30, № 3. С. 290–297. <https://doi.org/10.34216/1998-0817-2024-30-3-290-297>

Research Article

THE NEED FOR LEGAL REGULATION OF THE “BLACK BOX” OF ARTIFICIAL INTELLIGENCE

Kirill S. Golovin, Candidate of Jural Sciences, Institute of Legislation and Comparative Law under the Government of the Russian Federation, attorney at Moscow Region Bar Association. Moscow, Russia, golovinks@gmail.com, <https://orcid.org/0009-0002-8844-488X>

Annotation. This paper discusses the transparency and regulation of Artificial Intelligence (hereinafter the AI). It discusses the growing trend of using decision-making algorithms in various areas of life and the need for transparent and the understandable AI systems. The author highlights the concept of a “black box” where algorithms and their processes become hidden to users and developers, and emphasises the importance of explainability and accessibility of information about the operation of the AI. The article also addresses the right to explanation and possible strategies for ensuring transparency and accountability in algorithmic decision-making, including parameter documentation and system certification.

Keywords: artificial intelligence (AI) transparency, AI black box, AI regulation, algorithmic decision making, AI legal aspects, AI systems’ certification, challenging AI decisions.

For citation: Golovin K.S. The need for legal regulation of the “black box” of artificial intelligence. Vestnik of Kostroma State University, 2024, vol. 30, no. 3, pp. 290–297 (In Russ.). <https://doi.org/10.34216/1998-0817-2024-30-3-290-297>

Распространение алгоритмического принятия решений в повседневной жизни – это растущее явление, напрямую затрагивающее множество людей. Искусственный интеллект (ИИ) все больше проникает в различные виды человеческой деятельности, что приводит к принятию решений, зачастую выходящих за рамки понимания обычных людей. Этот процесс требует создания прозрачных и понятных систем ИИ, которые помогут пользователям лучше понимать, как принимаются решения.

Прозрачность ИИ – это не просто техническая необходимость, но и фундаментальный принцип, который должен управлять разработкой, внедрением и эксплуатацией систем ИИ. Принципы Национальной стратегии развития искусственного интеллекта в Российской Федерации до 2030 года¹, а также международные нормативные акты, такие как GDPR², подчеркивают важность прозрачности как средства защиты прав и свобод человека. Прозрачные системы ИИ способствуют защите прав пользователей и повышают доверие к таким технологиям [Goodman, Flaxman: 51].

На сегодняшний день системы ИИ регулируются весьма ограниченно: предполагается, что ИИ будут использоваться для помощи в принятии решений, а окончательное решение будет приниматься человеком на основе здравого смысла. Однако с ростом использования ИИ в более значимых и распространенных областях возрастает необходимость усиления регулирования. Важно, чтобы правовые рамки адаптировались к новым вызовам, связанным с использованием ИИ.

В правовом контексте технологий искусственного интеллекта термин «черный ящик» обретает особую важность. Традиционное развитие ИИ происходило под влиянием человеческого опыта и знаний, что обеспечивало прозрачность и предсказуемость его решений. Однако с появлением машинного обучения, особенно глубокого обучения, эта прозрачность существенно снизилась, уступив место высокой эффективности и способности к самообучению. В результате возникло понятие «черного ящика» – систем, внутренние процессы и алгоритмы которых остаются непонятными как для пользователей, так и для разработчиков [Pasquale: 10].

Законодательное определение и регулирование «черного ящика» становится особенно важно при правовом регулировании ИИ. Такие системы часто работают на основе сложных алгоритмов и данных, которые трудно интерпретировать. Это создает проблемы для всех сторон, так как невозможно понять, как и почему система принимает решения. Это непонимание может привести к непредсказуемым и негативным последствиям, особенно в таких критически важных областях, как здравоохранение, правосудие и финансы.

Все это подчеркивает важность определения правового статуса ИИ. В правовой науке существует множество точек зрения на этот вопрос. Одни авторы считают, что ИИ следует признать субъектом права, другие – что он обладает элементами правосубъектности, а третьи рассматривают его как объект права [Конвергенция: 142]. Однако признание ИИ субъектом права требует от него наличия правоспособности и дееспособности, что в текущих правовых рамках невозможно без понимания принципов работы ИИ и возможности интерпретации его решений.

А.А. Васильев и Ю.В. Печатнова определяют ИИ как «специализированный программно-вычислительный комплекс с встроенной способностью к самообучению, позволяющий имитировать когнитивные функции человека и автономно выполнять задачи» [Васильев, Печатнова: 16]. А.В. Тюлин рассматривает ИИ как систему без эмоций и чувств, не обладающую правосубъектностью, но способную имитировать когнитивные функции человека [Тюлин: 81]. А. Гурко утверждает, что для понимания проблем необходимо определить ИИ как науку и технологию создания интеллектуальных машин, выполняющих творческие функции, традиционно присущие человеку [Гурко: 5]. И.В. Понкин и А.И. Редькина описывают ИИ как систему, способную к самореферентной адаптации (способности системы адаптироваться на основе собственного опыта), глубинному самообучению и выполнению антропоморфных когнитивных функций [Понкин, Редькина: 95].

Кроме того, нормативное определение содержится в Указе Президента Российской Федерации «О развитии искусственного интеллекта в Российской Федерации», где ИИ определяется как совокупность технологических решений, имитирующих когнитивные функции человека и достигающих результатов, сравнимых с интеллектуальной деятельностью человека. Этот комплекс включает информационно-коммуникационную инфраструктуру и программное обеспечение (включая методы машинного обучения). Он также охватывает процессы и сервисы, связанные с обработкой данных и поиском решений. Федеральный закон от 24.04.2020 № 123-ФЗ также включает понятия ИИ-технологий (технологии искусственного интеллекта), такие как компьютерное зрение и обработка естественного языка.³

Данные определения не решают вопросов контроля ИИ, ответственности за принятые решения, а только соотносят его с человеком или его функциями. В тоже время если мы хотим допустить ИИ до принятия решений, которые будут влиять на жизни людей, необходим четкий контроль и понимание, как тот или иной механизм работает. На ИИ влияют разработчики, пользователи и тренеры, поэтому было бы неправильно говорить, что сам ИИ может

нести ответственность за свои действия, так как имеет свойства, похожие на человеческие [Головин: 268].

Как отмечает Д.Д. Гримм, цель научного исследования заключается в объяснении сложных явлений путем их упрощения [Гримм: 218]. Таким образом, важным является разработка нормативного определения важнейшего элемента-свойства в понимании ИИ – «черного ящика» и его регулирования. Это позволит обеспечить прозрачность и подотчетность ИИ-технологий. Для этого требуется:

1. *Определение термина «черный ящик»:* необходимо формализовать, что подразумевается под этим понятием в контексте ИИ. Это может включать критерии, такие как сложность алгоритмов и недоступность их интерпретации, по которым система признается «черным ящиком»;

2. *Регулирование разработки и использования ИИ.* Помимо определения термина «черный ящик», необходимо также регулировать разработку и использование ИИ. Законы и нормативные акты должны требовать от разработчиков предоставления информации о принципах работы их систем, особенно если они относятся к категории «черных ящиков». Это может включать требования по документации, прозрачности алгоритмов и отчетности.

3. *Механизмы проверки и сертификации.* Создание независимых органов или комиссий, которые будут проверять и сертифицировать ИИ-системы, особенно те, которые могут классифицироваться как «черные ящики». Это позволит обеспечить, что такие системы соответствуют установленным стандартам и не представляют угрозы для пользователей.

4. *Правовые последствия и ответственность.* Установление правовой ответственности (возможности привлечения к юридической ответственности) за использование неконтролируемого ИИ с «черным ящиком». Это включает определение субъектов, ответственных за ошибки или злоупотребления, возникающие вследствие использования таких систем.

Без этих мер правовая система не сможет адекватно регулировать использование ИИ. Таким образом, важно разработать нормативные определения и механизмы регулирования, которые позволят обеспечить прозрачность и подотчетность ИИ-технологий.

Под ИИ – «черным ящиком» понимается любой инструмент, использующий методы обработки естественного языка, машинного обучения или текстового анализа, который оперирует данными, недоступными для субъектов данных, и/или использует алгоритмы, которые также недоступны или настолько сложны, что их невозможно свести к понятной для субъекта данных серии правил.

Некоторые нормативно-правовые акты, регулирующие функционирование автоматизированных систем принятия решений, аналогичных указанным

выше, содержатся в законодательных положениях, касающихся обработки персональных данных. В соответствии со статьей 16 Федерального закона «О персональных данных»⁴ в Российской Федерации принятие решений, имеющих юридические последствия, исключительно на основе автоматизированной обработки персональных данных запрещено. Исключением являются ситуации, когда такие решения принимаются на основании письменного согласия субъекта данных либо в соответствии с положениями, установленными федеральным законодательством.

Следует отметить, что действующее законодательство о персональных данных не предоставляет исчерпывающего перечня ситуаций, в которых допускается использование автоматизированных систем принятия решений. Для обеспечения более четкого правового регулирования применения алгоритмов следует использовать терминологию, однозначно определяющую технологические аспекты принятия юридически значимых решений. В тексте Федерального закона «О персональных данных» используется термин «автоматизированная обработка персональных данных», который подразумевает использование компьютерных технологий для обработки данных. Однако этот термин охватывает лишь процессуальную сторону, а не технологическую, что создает определенные проблемы, поскольку алгоритмы представляют собой именно технологию, требующую специализированного правового регулирования [Алгоритмическая прозрачность: 146].

Термин «автоматизированная обработка персональных данных» ограничивается исключительно персональными данными, тогда как автоматизированные системы принятия решений могут обрабатывать и иные виды данных. Более того, данный термин не предполагает принятия окончательного решения, что может вводить в заблуждение, создавая впечатление, будто решение принимает человек.

Для обеспечения более четкого регулирования применения алгоритмов необходимо оперировать термином, однозначно определяющим технологию принятия юридически значимых решений. В Федеральном законе «О персональных данных» используется понятие «автоматизированная обработка персональных данных», что подразумевает использование вычислительной техники для обработки данных. Однако этот термин описывает процесс, а не технологию, что создает определенные трудности, так как алгоритмы представляют собой технологию, требующую правового регулирования.

Термин «автоматизированная обработка персональных данных» охватывает только персональные данные, в то время как автоматизированные системы принятия решений анализируют и другие виды данных. Вдобавок данный термин не предполагает

принятия конечного решения, что может ввести в заблуждение, создавая впечатление, что решение принимает человек.

Пользователи и участвующие стороны (например, потерпевшие в случае причинения вреда) должны иметь возможность понять, как было принято то или иное решение и на каких данных оно базировалось. Такой подход способствует не только доверию и принятию технологий ИИ, но и облегчает идентификацию и корректировку ошибок или предвзятостей в алгоритмах [Burrell: 10].

Многие современные опасения относительно ИИ – «черного ящика» возникают из-за технологического феномена, согласно которому ИИ часто представляет собой вид машинного обучения, включая «глубокие нейронные сети», которые увеличивают непрозрачность. Эти системы ИИ не воспроизводят пошаговую логику и правила, которые могли бы быть интуитивно понятны. Вместо этого они выявляют закономерности из огромных наборов примеров для создания всё более точных прогнозов. С ростом использования таких инструментов возрастают и наши опасения по поводу их непрозрачной природы.

С самого начала возникает вопрос о том, является ли право на объяснение отдельным (*per se*) правом субъекта правоотношений или это лишь элемент более широкого права – права на информацию. Некоторые ученые ставят под сомнение само существование такого права [Wachter, Mittelstadt, Floridi: 5], в то время как другие отмечают, что, независимо от того, как определяется право на объяснение, оно не является «иллюзорным» [Selbst, Powles: 235]. Несомненно, право на объяснение служит конкретной цели – дать человеку возможность оспорить правильность решения, принятого алгоритмом. Как указывает С.В. Мельник, без понимания того, на каких критериях и факторах было основано решение, это право не может быть реализовано на практике [Мельник: 20].

Алгоритмы принятия решений в ИИ используют исторические данные для построения прогностических моделей. Таким образом, объяснения требуют именно модель, а не алгоритм как таковой, поскольку разные алгоритмы могут давать схожие модели.

Без понимания и регулирования «черного ящика» ИИ невозможно обеспечить справедливое правосудие с участием ИИ. Необходимо четко формализовать понятие «черного ящика», устанавливая критерии, по которым система признается таковой, и требовать от разработчиков предоставления информации о принципах работы их систем. Независимые органы должны проверять и сертифицировать такие системы, чтобы они соответствовали установленным стандартам и не представляли угрозы. Также следует установить правовую ответственность за ошибки

или злоупотребления, связанные с использованием «черных ящиков».

Как указывает А.А. Антонов, к правосудию с использованием достижений искусственного интеллекта следует относиться с осторожностью, так как правосудие осуществляется вследствие правового спора и необходимости его справедливого разрешения. При этом возникают сомнения в возможности ИИ принимать правильные решения [Антонов: 70].

Разумное поведение человека сочетает в себе рациональные и эмоциональные составляющие, учитывающая правила морали, этики и нравственности. Разумность деятельности определяется сознанием человека, что позволяет адекватно мыслить, переживать события и осознавать свои действия. Поэтому для обеспечения правосудия с участием ИИ важно, чтобы объяснения решений алгоритмов соответствовали человеческим рассуждениям и социальным нормам. Это подчеркивает необходимость нормативного определения и регулирования «черного ящика», что обеспечит прозрачность и подотчетность ИИ-технологий в судебной системе [Кляшторный: 108].

Право на объяснение активируется, когда результат работы автоматизированного инструмента отличается от ожиданий пользователя. Это побуждает его искать контрастное объяснение, выясняя причины конкретного результата и возможные изменения для достижения ожидаемого результата. В судебном контексте, где алгоритмы помогают принимать решения, объяснения должны соответствовать человеческим рассуждениям и социальным нормам, преодолевая разрыв между техническими процессами и общественным пониманием.

Реализация подсистемы регистрации событий была определена архитекторами систем информационных технологий как жизнеспособная стратегия документирования значимых операционных параметров алгоритмов, позволяющая впоследствии восстановить их деятельность. Этот подход, схожий с использованием бортовых самописцев в авиации для записи данных о полете, предлагает обязательное документирование обоснований принятия решений алгоритмами машинного обучения. Этот метод, который Марго Камински называет «квалифицированной прозрачностью», призван обеспечить различным заинтересованным сторонам – от частных лиц и экспертов до регулирующих органов – индивидуальный доступ к информации, относящейся к решениям алгоритмов [Kaminski: 20].

Преимущества такого подхода заключаются в простоте реализации, незначительном влиянии на стоимость развертывания и обслуживания системы, а также в отсутствии необходимости в сложных процедурах проверки. Дорогостоящие нормативные требования могут исказить динамику рынка или вы-

нудить поставщиков переехать в менее регулируемые регионы.

Безопасность и целостность регистрируемых данных можно обеспечить с помощью криптографических мер, а управление данными можно делегировать специализированным поставщикам услуг или надежным третьим сторонам, что сведет к минимуму несанкционированные изменения данных. Кроме того, государственные органы могут контролировать хранение таких данных, что повышает безопасность конфиденциальной информации, например, в машинных системах кредитного скоринга, контролируемых финансовыми регуляторами. Эта система может привести к установлению требований к регистрации данных в конкретных секторах, что облегчит как решение индивидуальных проблем, связанных с точностью принятия решений, так и надзор за системными операциями со стороны регулирующих органов.

Однако это решение не является универсальным, особенно для сложных алгоритмов машинного обучения, таких как глубокие нейронные сети, где запутанная сеть весов и взаимодействий признаков не поддается простой интерпретации и документированию. Кроме того, динамическая природа систем машинного обучения, которые развиваются с появлением новых данных, усложняет процесс протоколирования. В отличие от традиционных ИТ-систем, где основной пользователь задает параметры записи данных, системы машинного обучения требуют саморегулирующегося механизма для принятия решений о том, что и когда регистрировать. Такое саморегулирование создает проблемы для поддержания прозрачности, поскольку перекладывает ответственность за определение критериев протоколирования с разработчиков на саму систему, что потенциально создает уязвимость для ошибок или манипуляций. Следовательно, это требует формального пересмотра самого процесса регистрации событий, по сути заменяя поиск алгоритмической прозрачности задачей обеспечения прозрачности самого механизма регистрации.

Критической проблемой для эффективности системы регистрации событий является ее неспособность полностью учесть контекст, в котором принимаются алгоритмические решения. На функциональность алгоритма влияют не только входные данные и механизмы их обработки (которые могут быть относительно просто задокументированы), но и исторический контекст принятия решений. Чтобы полностью понять обоснование текущих результатов работы алгоритма, может потребоваться изучение обширного хранилища данных с подробным описанием его предыдущих решений. Без доступа к такому всестороннему историческому контексту простое документирование текущих параметров, используемых при принятии решений, может оказаться недостаточным для реконструкции

и оценки работы алгоритма на предмет корректности и справедливости. Эта проблема становится все более очевидной при использовании алгоритмов, в значительной степени основанных на машинном обучении, что усложняет задачу достижения прозрачности системы с помощью протоколирования.

Еще одно существенное препятствие – расшифровка связи между записанными параметрами и внутренней логикой алгоритма. Даже если удастся преодолеть первоначальные препятствия и точно задокументировать ключевые параметры для воспроизведения процесса принятия решений алгоритмом, доступ к основной логике алгоритма остается сложной задачей. Учитывая защиту систем в соответствии с законами об интеллектуальной собственности, отсутствие доступа к исходному коду алгоритма препятствует всестороннему пониманию логики принятия решений, необходимой для оценки его корректности [Gervais: 117].

Кроме того, эффективность регистрации событий в достижении поставленной цели требует тщательного изучения. Хотя сторонники прозрачности стремятся повысить надежность алгоритмических операций, одной прозрачности может быть недостаточно для обеспечения надежности обработки или защиты прав личности. Истинная справедливость алгоритмической обработки требует не только проверки точности работы, но и соблюдения правовых и этических норм. Следовательно, ИТ-система, точно определяющая взаимосвязи данных, все равно не может считаться надежной, если она не отвечает этим более широким критериям. Если полагаться только на регистрацию параметров обработки, это не позволит выявить недостатки, связанные с внешними данными, на которых основан алгоритм, что подчеркивает ограниченность такого подхода в достижении комплексной прозрачности и надежности системы.

Удостоверение алгоритмических операций с помощью сертификации представляет собой альтернативный подход, сосредоточенный на формальной оценке всей системы, а не на попытках отследить правильность отдельных решений. Этот подход предполагает создание национальной или международной системы сертификации, специально предназначенной для систем машинного обучения. Цель такой системы – не только обеспечить соответствие установленным нормам и стандартам на этапах проектирования, создания и тестирования автоматизированных систем принятия решений. Она также будет направлена на статистическую проверку надежности решений, принимаемых этими системами.

Внедрение системы сертификации может решить несколько ключевых проблем, предлагая технически жизнеспособное решение, которое не требует специальных знаний от людей, принимающих автоматиче-

ские решения. Однако внедрение системы сертификации требует решения различных проблем, включая разнообразие существующих механизмов сертификации в ИТ-индустрии. Эти механизмы, как правило, делятся на две категории: механизмы, проверяющие процессы разработки и сопровождения программного обеспечения (сертификация процессов), и механизмы, удостоверяющие подлинность и целостность программного обеспечения (сертификация кода), каждый из которых регулируется отдельными нормами и стандартами.

Выводы

Современное развитие искусственного интеллекта (ИИ) и его интеграция в различные сферы жизни вызывает необходимость создания новых правовых рамок, направленных на обеспечение прозрачности, подотчетности и справедливости в алгоритмических решениях.

Гражданам и юридическим лицам должно быть предоставлено право на объяснение решений, принимаемых с использованием ИИ, что позволит оспаривать такие решения и защищать свои права.

Для повышения прозрачности следует ввести обязательное документирование ключевых параметров, использованных при принятии алгоритмических решений. Это позволит восстановить и проанализировать действия ИИ, обеспечивая возможность для их проверки и коррекции.

Необходима разработка национальной или международной системы сертификации для алгоритмических систем, которая будет оценивать их надежность и соответствие установленным стандартам на всех этапах жизненного цикла.

Внедрение подхода с использованием системы регистрации событий, аналогичной бортовым самописцам в авиации, позволит документировать обоснования решений ИИ и предоставлять эту информацию заинтересованным сторонам.

Обеспечение безопасности и целостности данных, используемых ИИ, через криптографические методы и управление данными специализированными службами или доверенными третьими сторонами.

Для реализации вышеуказанных выводов предлагаю следующие формулировки статьи:

Статья 1261.1 ГК РФ «Искусственный интеллект»

1. Каждый пользователь имеет право на объяснение решения, принятого с использованием системы искусственного интеллекта, если это решение затрагивает его права и свободы. Такое объяснение должно включать в себя информацию о данных, использованных для принятия решения, и логику алгоритма.

2. Все системы искусственного интеллекта, используемые для принятия решений, обязаны документировать ключевые параметры, используемые в процессе принятия решения, включая значимые

операционные параметры и обоснования принятых решений. Данные должны храниться в зашифрованном виде и быть доступны для анализа заинтересованными сторонами и регулируемыми органами.

3. Обязательная сертификация систем искусственного интеллекта, предназначенных для принятия решений, затрагивающих права и свободы. Сертификация проводится уполномоченными органами и подтверждает соответствие системы установленным стандартам безопасности, надежности и прозрачности.

4. Все автоматизированные системы принятия решений обязаны регистрировать значимые операционные параметры и обоснования принятых решений. Данные регистрации должны храниться в зашифрованном виде и быть доступны для анализа заинтересованными сторонами и регулируемыми органами.

Эти изменения в гражданском законодательстве помогут обеспечить прозрачность, подотчетность и справедливость в использовании ИИ, защищая права граждан и способствуя развитию надежных и безопасных технологий.

Примечания

¹ Указ Президента Российской Федерации от 10 октября 2019 г. № 490 «О развитии искусственного интеллекта в Российской Федерации» (Национальная стратегия развития искусственного интеллекта на период до 2030 года). URL: <http://www.kremlin.ru/acts/bank/44731>.

² General Data Protection Regulation. URL: <https://eur-lex.europa.eu/eli/reg/2016/679/oj> (дата обращения: 25.06.2024).

³ Федеральный закон от 24.04.2020 № 123-ФЗ «О проведении эксперимента по установлению специального регулирования в целях создания необходимых условий для разработки и внедрения технологий искусственного интеллекта в субъекте Российской Федерации – городе федерального значения Москве и внесении изменений в статьи 6 и 10 Федерального закона «О персональных данных» // СЗ РФ. 27.04.2020. № 17. Ст. 2701.

⁴ Федеральный закон от 27 июля 2006 г. № 152-ФЗ «О персональных данных» (ред. от 24.04.2020) // СЗ РФ. 2006. № 31 (ч. 1). Ст. 3451.

Список литературы

Антонов А.А. Искусственный интеллект как источник повышенной опасности // Юрист. 2020. № 7. С. 69–74.

Васильев А.А., Печатнова Ю.В. Искусственный интеллект, право, проблемы и перспективы // Российско-Азиатский правовой журнал. 2020. № 2. С. 14–18.

Головин К.С. Предложения по реформированию гражданско-правовой охраны искусственного

интеллекта как единого объекта интеллектуальной собственности // Актуальные проблемы государства и права. 2023. Т. 7, № 2 (26). С. 263–274.

Гримм Д.Д. Основы учения о юридической сделке в современной немецкой доктрине пандектного права. Прологомены к общей теории гражданского права. Т. 1 (часть вторая) // Вестник гражданского права. 2015. № 2. С. 95.

Гурко А. Искусственный интеллект и авторское право: взгляд в будущее // ИС. Авторское право и смежные права. 2017. № 12. С. 7–18.

Кляшторный А.С. Критическое осмысление онтологических оснований концепции «сильного» искусственного интеллекта в контексте проблематики теории действия Х. Йоаса // Уральская аналитическая конференция. Екатеринбург, 2014. С. 106–117.

Конвергенция частноправового регулирования общественных отношений сквозь призму эффективности права: монография / К.А. Абдуллаев, А.В. Абрамова, Э.А. Абызова и др.; отв. ред. А.Н. Левушкин, Э.Х. Надышева. Москва: Юстицинформ, 2023. 672 с.

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. № 6 (163). С. 146.

Мельник С.В. Основные подходы к определению формы защиты права в отечественной цивилистической доктрине // Актуальные вопросы науки и практики: сб. науч. ст. Орел, 2010. С. 20.

Понкин И.В., Редькина А.И. Искусственный интеллект с точки зрения права // Вестник Российского университета дружбы народов. Сер.: Юридические науки. 2018. Т. 22, № 1. С. 93.

Тюлин А.В. Трансформация правовой реальности, вызванная системами искусственного интеллекта // Актуальные проблемы эффективности частного права: монография / отв. ред. А.Н. Левушкин. Москва: Юстицинформ, 2022. 464 с.

References

Antonov A.A. *Iskusstvennyi intellekt kak istochnik povyshennoi opasnosti* [Artificial Intelligence as a Source of Increased Danger]. *Yurist* [Lawyer], 2020, no. 7, pp. 69-74. (In Russ.)

Golovin K.S. *Predlozheniya po reformirovaniyu grazhdansko-pravovoi okhrany iskusstvennogo intellekta kak edinogo obekta intellektualnoi sobstvennosti* [Proposals for the Reform of Civil Law Protection of Artificial Intelligence as a Single Object of Intellectual Property]. *Aktualnye problemy gosudarstva i prava* [Actual problems of the state and law], 2023, vol. 7, no. 2 (26), pp. 263-274. (In Russ.)

Grimm D.D. *Osnovy ucheniya o yuridicheskoi sdelke v sovremennoi nemetskoj doktrine pandektnogo prava. Prolegomeny k obshchei teorii grazhdanskogo prava*. T. 1 (chast vtoraya) [Fundamentals of the Doctrine of Legal Transaction in Modern German Pandect Law Doctrine. Prolegomena to the General Theory of Civil Law. Vol. 1 (part two)]. *Vestnik grazhdanskogo prava* [Civil Law Bulletin], 2015, no. 2, pp. 95. (In Russ.)

Gurko A. *Iskusstvennyi intellekt i avtorskoe pravo: vzglyad v budushchee* [Artificial Intelligence and Copyright: A Look into the Future]. *Avtorskoe pravo i smezhnyye prava* [Copyright and related rights], 2017, no. 12, pp. 7-18. (In Russ.)

Klyashtornyi A.S. *Kriticheskoe osmyslenie ontologicheskikh osnovanii kontseptsii "sil'nogo" iskusstvennogo intellekta v kontekste problematiki teorii deistviya* Kh. Yoyasa [Critical Reflection on the Ontological Foundations of the "Strong" Artificial Intelligence Concept in the Context of H. Joas's Theory of Action]. *Uralskaya analiticheskaya konferentsiya* [Ural analytical conference]. Yekaterinburg, 2014, pp. 106-117. (In Russ.)

Konvergentsiya chastnopravovogo regulirovaniya obshchestvennykh otnoshenii skvoz prizmu effektivnosti prava: monografiya [The Convergence of Private Law Regulation of Social Relations through the Prism of Legal Efficiency: monograph], K.A. Abdullaev, A.V. Abramova, E.A. Abyzova et al.; ed. by A.N. Levushkin, E.Kh. Nadyшева. Moscow, Yustitsinform Publ., 2023, 672 p. (In Russ.)

Kuteynikov D.L., Izhayev O.A., Zenin S.S., Lebedev V.A. *Algoritmicheskaya prozrachnost' i podotchetnost': pravovye podkhody k razresheniyu problemy "Chyornogo yashchika"* [Algorithmic Transparency and Accountability: Legal Approaches to Solving the "Black Box" Problem]. *Lex Russica*, 2020, no. 6 (163), pp. 146. (In Russ.)

Melnik S.V. *Osnovnye podkhody k opredeleniyu formy zashchity prava v otechestvennoy tsivilisticheskoi doktrine* [Main Approaches to Defining the Form of Legal Protection in Domestic Civil Law Doctrine]. *Aktualnye voprosy nauki i praktiki: sbornik nauchnykh statei* [Current issues of science and practice]. Orel, 2010, p. 20. (In Russ.)

Ponkin I.V., Redkina A.I. *Iskusstvennyi intellekt s tochki zreniya prava* [Artificial Intelligence from the Legal Perspective]. *Vestnik Rossiiskogo universiteta druzhby narodov. Ser.: Yuridicheskie nauki* [Bulletin of the Peoples' Friendship University of Russia. Ser.: Legal Sciences], 2018, vol. 22, no. 1, p. 93. (In Russ.)

Tiulin A.V. *Transformatsiya pravovoi realnosti, vyzvannaya sistemami iskusstvennogo intellekta* [Transformation of Legal Reality Caused by Artificial Intelligence Systems]. *Aktualnye problemy effektivnosti chastnogo prava: monografiya* [Current Problems of the Efficiency of Private Law: monograph], ed. by A.N. Levushkin. Moscow, Yustitsinform Publ., 2022, 464 p. (In Russ.)

Vasiliev A.A., Pechatnova Yu.V. *Iskusstvennyi intellekt, pravo, problemy i perspektivy* [Artificial Intelligence, Law, Problems and Prospects]. *Rossiysko-Azi-*

atskii pravovoi zhurnal [Russian-Asian Law Journal], 2020, no. 2, pp. 14-18. (In Russ.)

Burrell J. How the machine ‘thinks’: Understanding opacity in machine learning algorithms. *Big Data & Society*, 2016, vol. 3 (1). URL: <https://journals.sagepub.com/doi/epub/10.1177/2053951715622512> (дата обращения: 24.06.2024).

Daniel G. Is Intellectual Property Law Ready for Artificial Intelligence? *GRUR International*, 2020, vol. 69, iss. 2, pp. 117-118. <https://doi.org/10.1093/grurint/ikz025> (access date: 24.06.2024).

Goodman B., Flaxman S. European Union Regulations on Algorithmic Decision-Making and a “Right to Explanation”. *AI Magazine*, 2017, vol. 38 (3), pp. 50-57. <https://doi.org/10.1609/aimag.v38i3.2741>

Kaminski M.E. The Right to Explanation, Explained. U of Colorado Law Legal Studies Research Paper No. 18-24. *Berkeley Technology Law Journal*, 2019, vol. 34, no. 1. <http://dx.doi.org/10.2139/ssrn.3196985> (access date: 24.06.2024).

Pasquale F.A. *The Black Box Society: The Secret Algorithms that Control Money and Information*. Book Gallery, 2015, vol. 96. URL: <https://digitalcommons.law.umaryland.edu/books/96> (access date: 24.06.2024)

Selbst A.D., Powles J. Meaningful Information and the Right to Explanation. *International Data Privacy Law*, 2017, vol. 7 (4), pp. 233-242. URL: <https://ssrn.com/abstract=3039125> (access date: 24.06.2024).

Wachter S., Mittelstadt B., Floridi L. Transparent, Explainable, and Accountable AI for Robotics. *Science Robotics*, 2017, vol. 2. URL: https://discovery.ucl.ac.uk/id/eprint/10038294/1/Wachter_Transparent_explainable_accountable_AI.pdf (access date: 24.06.2024).

Статья поступила в редакцию 24.06.2024; одобрена после рецензирования 02.08.2024; принята к публикации 03.09.2024.

The article was submitted 24.06.2024; approved after reviewing 02.08.2024; accepted for publication 03.09.2024.