Дата поступления рукописи в редакцию: 28.03.2026 г.
Дата принятия рукописи в печать: 03.05.2026 г.
DOI: 10.24412/2782-3830-2026-4-496-504

ЧУМАЧЕНКО Василий Андреевич,
3 курс аспирантуры Санкт-Петербургский Государственный Университет,
e-mail: mail@law-books.ru

СОЗДАНИЕ И ОБУЧЕНИЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ЦЕЛЯХ ЗАЩИТЫ ПУБЛИЧНОГО ИНТЕРЕСА: ЗА И ПРОТИВ

Аннотация. *В статье проводится правовой анализ применения искусственного интеллекта (далее - ИИ) для защиты публичного интереса. В отличие от распространенного подхода, фокусирующегося на этических дилеммах, статья переводит проблему в плоскость публично-правовых ограничений, возникающих на этапе обучения ИИ. Через призму жизненного цикла данных рассматриваются конституционные гарантии приватности, запреты на использование отдельных категорий информации и проблема соблюдения авторских прав при формировании датасетов. На основе законодательства и актов ЕС формулируются процедурные требования к созданию и обучению ИИ для государственных нужд: фильтрация датасетов, риск-ориентированный подход, прозрачность источников. Автор приходит к выводу, что реализация подобных систем требует формирования специальной правовой доктрины, основанной на принципах человеческого контроля, прозрачности, распределения ответственности и публичного надзора на всех стадиях обработки данных. Только при этих условиях ИИ может стать легитимным инструментом, усиливающим, а не подменяющим правовые и демократические основы общества.*

Ключевые слова: *государство, безопасность, искусственный интеллект, общество, публичный интерес, обучение, алгоритм, прозрачность.*

CHUMACHENKO Vasily Andreevich,
3rd year of postgraduate studies at St. Petersburg State University

CREATION AND TRAINING OF ARTIFICIAL INTELLIGENCE IN ORDER TO PROTECT THE PUBLIC INTEREST: PROS AND CONS

Annotation. *The article provides a legal analysis of the use of artificial intelligence (AI) to protect the public interest. In contrast to the common approach that focuses on ethical dilemmas, the article translates the problem into the plane of public law restrictions that arise at the stage of AI training. Constitutional guarantees of privacy, prohibitions on the use of certain categories of information, and the problem of copyright compliance in the formation of datasets are considered through the prism of the data lifecycle. Based on EU legislation and acts, procedural requirements for the creation and training of AI for government needs are formulated: dataset filtering, risk-based approach, transparency of sources. The author concludes that the implementation of such systems requires the formation of a special legal doctrine based on the principles of human control, transparency, responsibility sharing and public oversight at all stages of data processing. Only under these conditions can AI become a legitimate tool that strengthens, rather than replaces, the legal and democratic foundations of society.*

Key words: *state, security, artificial intelligence, society, public interest, learning, algorithm, transparency.*

Актуальность рассматриваемой темы обусловлена стремительным процессом цифровизации государственного управления, который в настоящее время охватывает все сферы жизни общества. В рамках этого процесса система искусственного интеллекта (далее – ИИ) начинает занимать центральное место в обеспечении национальной безопасности, поддержании общественного порядка и защите ключевых национальных интересов. Российская Федера-

ция, следуя общемировым трендам, активно занимается разработкой и внедрением различных решений на основе ИИ. Эти технологии направлены на прогнозирование правонарушений, автоматизацию процессов модерации в информационном пространстве и повышение эффективности работы государственных органов. Национальная стратегия развития искусственного интеллекта, которая была установлена на период до 2030 года, четко обозначает приоритетное использование ИИ в интересах как общества, так и государства. Это требует создания мощных алгоритмических моделей, которые должны быть обучены на обширных и разнообразных массивах данных [17].

Введение искусственного интеллекта в область права и публичного управления является одним из самых значительных и спорных вызовов нашего времени. Аргументы в пользу интеграции известны: повышение эффективности и объективности государственных институтов, снижение влияния человеческого фактора и коррупционных рисков. Так, ИИ способен выявлять скрытые нарушения в госзакупках, оптимизировать распределение ресурсов в здравоохранении или прогнозировать чрезвычайные ситуации. Вместе с тем, существуют весомые контраргументы, которые подчеркивают серьезные риски, угрожающие основам правопорядка и демократическим ценностям, связанные с проблемой «черного ящика», алгоритмической дискриминации и подрыва доверия к правовой системе.

Развитие искусственного интеллекта в России получает поддержку на самом высоком уровне — со стороны руководства страны. В частности, Президент Российской Федерации подписал Указ №124 от 15 февраля 2024 года, вносящий существенные перемены в Общегосударственный план развития искусственного интеллекта, действующий до 2030 года. Эти перемены важны и достойны особого внимания, так как затрагивают коренные стороны становления этой отрасли в стране. План намечает смелые цели, главные из которых включают рост благосостояния народа, улучшение качества жизни граждан, укрепление безопасности державы и поддержание правопорядка [3].

Для безопасного и эффективного развития и внедрения технологий искусственного интеллекта необходимо создать новую регуляторную среду, которая будет учитывать и балансировать интересы всех вовлеченных сторон: человека, общества, государства, а также организаций. занимающихся разработкой систем ИИ и робототехники, и потребителей, которые используют эти технологии и услуги. Данная концепция согласуется с общепринятой идеей о «публичном интересе», который определяется как интерес, закрепленный в законодательстве, обеспечиваемый и охраняемый правом. Публичный интерес в цифровую эпоху выступает как сложный, динамичный баланс между коллективной безопасностью, эффективностью и защитой фундаментальных прав каждой личности. Важно чтобы развитие технологий ИИ не только соответствовало современным требованиям, но и учитывало эти потребности, обеспечивая тем самым защиту прав граждан и интересов общества [5].

Основное преимущество таких систем заключается в их умении выявлять скрытые взаимосвязи и принимать решения в условиях неопределенности, не опираясь на жесткие алгоритмические предписания. В узкопрофильных задачах эффективность ИИ сопоставима с человеческим интеллектом, однако при анализе больших объемов данных компьютерные системы демонстрируют значительное превосходство [15].

Технологический фундамент ИИ формируют несколько взаимосвязанных компонентов. Информационно-коммуникационная инфраструктура, включающая высокопроизводительные вычислительные мощности и центры обработки данных, обеспечивает необходимую среду функционирования сложных моделей. Программное обеспечение (ядром его выступают методы машинного обучения) позволяет системам обучаться непосредственно на данных, выявлять паттерны, делать выводы. Функциональную основу составляют такие технологии: компьютерное зрение (с целью проведения анализа и интерпретации визуальной информации); обработка естественного языка для понимания, генерации человеческой речи; системы распознавания, синтеза голоса (с целью создания интуитивных интерфейсов взаимодействия). Интеллектуальная поддержка принятия решений помогает специалистам делать более обоснованные выводы в сложных управленческих или диагностических ситуациях.

Существующие правовые определения искусственного интеллекта дополняются множеством доктринальных подходов. Каждый из них акцентирует внимание непосредственно на различных аспектах ИИ как объекта правового анализа. Исследователи И.С. Шиткина и Д.О. Бирюков определяют ИИ как результат интеллектуальной деятельности, выраженный в наделении материальных объектов функциями, свойственными человеческому интеллекту [16]. Такой подход подчеркивает антропоцентричную природу технологии, ее происхождение из человеческого творчества и инженерной мысли. Альтернативная концепция рассматривает ИИ как программно-техническое средство, обладающее биоподобными способностями [10]. Эта трактовка указывает на его сложную технологическую сущ-

ность и способность имитировать поведенческие и мыслительные процессы живых организмов. Научные дискуссии подтверждают, что ИИ представляет собой не просто статичную программу, а сложную адаптивную систему, способную к самообучению.

Национальная стратегия в области ИИ формализует это понимание, выделяя два основных типа систем: слабый и сильный искусственный интеллект. «Слабый» (узкий) ИИ предназначен для решения конкретных, четко очерченных задач, таких как распознавание изображений, управление беспилотным транспортом или перевод текстов. Он не обладает сознанием или способностью к самостоятельному мышлению. «Сильный» (общий) ИИ — гипотетическая система, потенциал которой сопоставим с человеческим интеллектом в его полноте, включая способность мыслить абстрактно, взаимодействовать с миром и адаптироваться к любым новым условиям.

Невозможность дать однозначный ответ о правовой природе ИИ обусловлена, с одной стороны, многообразием мнений в российской и зарубежной юриспруденции, а с другой — его научно-технической сущностью. Поскольку искусственный интеллект является в первую очередь постоянно эволюционирующей технологической категорией, любое юридическое определение не может быть сформулировано в отрыве от этого контекста. Ключевой же вопрос о признании ИИ субъектом или объектом права напрямую зависит от фундаментальной проблемы определения самого интеллекта — темы междисциплинарных дискуссий между юристами, кибернетиками, нейробиологами и философами. Следовательно, пока междисциплинарный консенсус по этому вопросу не будет достигнут, формирование устойчивой правовой концепции остается невозможным.

Определения, предложенные в законе и доктрине, акцентируют внимание на таких ключевых характеристиках, как автономность и обучаемость – основополагающие свойства для понимания роли ИИ в современном обществе. Этап обучения ИИ имеет самостоятельное юридическое значение, так как в процессе обучения алгоритм анализирует большие объемы данных, представляющие собой обучающие выборки, конечным результатом такой деятельности становится выявление закономерностей и формирование внутренних моделей, которые потом используются для принятия решений [16]. Важно отметить, что именно на этапе обучения в систему могут быть заложены потенциальные предвзятости, ошибки и этические установки, что может иметь серьезные последствия. Качество и природа данных, на которых происходит обучение ИИ, напрямую влияют на правовые последствия его применения. Это означает, что процесс обучения должен рассматриваться не просто как техническая операция, а как действие, имеющее юридическое значение. Поэтому необходимо наличие собственных правовых оснований и гарантий, чтобы обеспечить корректное и этичное использование искусственного интеллекта [9].

Этап обучения алгоритмов является критически с точки зрения как эффективности, так и легитимности будущих решений ИИ. Когда речь идет о защите публичного интереса, например, в сферах национальной безопасности, борьбы с киберпреступностью или защиты детей, государство сталкивается с необходимостью обучать на особо чувственных данных, что порождает острые правовые коллизии.

Во-первых, это проблема обучения ИИ для выявления запрещенного контента, такого как детская порнография, материалы, пропагандирующие экстремизма или насилие. Для создания эффективных моделей распознавания необходимы обширные обучающие выборки, содержащие образцы такого контента. Это ставит вопрос о законности хранения и обработки самим государством или его контрагентами материалов, оборот которых строго запрещен. Возникает парадоксальная ситуация, при которой для борьбы с преступностью может создаваться де-факто «легальный оборот» криминального контента в исследовательских или технологических целях. Разрешение этой коллизии требует четкой правовой основы, определяющей строгие условия доступа, хранения, обезличивания и уничтожения таких данных, а также механизмы независимого контроля за этими процессами.

Во-вторых, серьезную озабоченность вызывает использование для обучения ИИ данных, защищенных профессиональной или личной тайной. Речь идет о применении в целях предиктивной аналитики или обеспечения безопасности перехваченной переписки, метаданных телекоммуникационного трафика, закрытых государственных и ли коммерческих баз данных. Ключевыми становятся вопросы о достаточности существующих правовых оснований, об обязательности предварительного судебного контроля, о необходимости уведомления граждан об использовании их данных в подобных целях, и о соразмерности таких мер провозглашаемым целям защиты публичного интереса.

Таким образом, сегодня острота дискуссии смещается из плоскости общих выгод/рисков в область конкретных юридических конструкций, центральной из которых является обучение ИИ как стадия обработки информации, подпадающая под действие публично-правовых ограниче-

ний. Для понимания правовой природы ИИ важно отметить, что действующее законодательство ориентировано на урегулирование «слабого» ИИ, который рассматривается как объект права [6].

Обучение искусственного интеллекта – это стадия жизненного цикла, на которой обращение с информацией приобретает самостоятельное юридическое значение. В этот момент технологический процесс пересекается с конституционными гарантиями и законодательными запретами. Так, статья 24 Конституции РФ устанавливает императивный запрет на сбор, хранение, использование и распространение информации о частной жизни лица без его согласия – это фундаментальное ограничение, которое не может быть преодолено ссылками на технологическую необходимость или публичную пользу алгоритма. Этот конституционный принцип получает свое развитие и детализацию в отраслевом законодательстве, прежде всего в Федеральном законе «О персональных данных». В соответствии с частью 3 статьи 3 Закона под обработкой персональных данных понимается любое действие (операция) или совокупность действий, совершаемых с использованием средств автоматизации или без их использования, включая сбор, запись, систематизацию, накопление, хранение, уточнение, извлечение, использование, передачу, обезличиванием, блокирование. Удаление, уничтожение персональных данных. Данный перечень функционально полностью совпадает с технологическими операциями, осуществляемыми при подготовке датасетов и обучении искусственного интеллекта: от сбора и разметки данных до формирования обучающих выборок и непосредственно машинного обучения. Отсюда вытекает ключевое правовое последствие: если в состав обучающего датасета входят персональные данные, вся совокупность операций по его созданию и использованию для обучения ИИ подпадает под действие Закона о персональных данных и требует наличия либо согласия субъекта персональных данных, либо иного законного основания, предусмотренного ч. 1-2 статьи 6 указанного Закона.

Важно подчеркнуть, что законодатель и стратегические документы в сфере развития искусственного интеллекта исходят из того же понимания. Так, Национальная стратегия развития ИИ на период до 2030 года прямо выделяет такие этапы, как «набор данных» и «разметка данных», подчеркивая при этом, что набор данных должен быть предварительно подготовлен с учетом требований законодательства Российской Федерации в области информации и защиты информации. Тем самым создается нормативный «мост» между технологическими стадиями жизненного цикла ИИ и правовыми требованиями: государство признает, что работа с данными не является технической вольностью, а должна осуществляться в строгих правовых рамках, установленных для оборота информации, особенно если она затрагивает конституционные права граждан.

В правовом поле формируется принципиально важный вывод: юридически значимыми становятся не только конечные решения, принятые системой ИИ, но и весь контур работы с информацией на этапе обучения. Именно на этой стадии закладываются как потенциальные риски для прав и свобод человека, так и возможности для их предотвращения, что требует самостоятельного правового регулирования и контроля.

Логическим продолжением данного подхода является распространение правовых ограничений не только на персональные данные, но и на иные категории информации, оборот которых прямо запрещен либо ограничен в публичных интересах. Принцип здесь действует тот же, но с еще более строгой императивностью: если распространение определенных сведений запрещено законом, их включение в обучающие выборки недопустимо – независимо от «технической доступности» таких материалов (например, найденных в открытом интернете). Технологическая возможность собрать больше данных не отменяет нормативных запретов.

Опорные режимы запрещенной информации в данном контексте представлены несколькими ключевыми блоками законодательства.

Во-первых, это законодательство о противодействии экстремизму. Включение в датасет материалов из федерального списка экстремистских материалов недопустимо, так как сам факт их хранения и обработки может образовывать состав правонарушения, предусмотренного в статье 282 УК РФ. Государство, борющееся с экстремизмом, не может само создавать массивы таких данных в исследовательских и технологических целях без специального правового режима.

Во-вторых, это законодательство о защите детей от вредной информации, которое фиксирует категории информации, причиняющей вред здоровью раз витию детей или запрещенной для распространения среди несовершеннолетних.

Важно сказать, что мировые правопорядки начинают вырабатывать подходы к этой проблеме. Регламент ЕС об искусственном интеллекте вводит концепцию систем «высокого риска», к которым относятся ИИ, используемые в управленческой и правоохранительной деятельности. Регламент к неприемлемым относит и системы ИИ используемые в правоохранительных целях для профилировании и типологизации людей, биометрической их категоризации с

целью установления их политических взглядов, религиозных предпочтений и т. д. Запрещен ИИ в сфере нецелевого сбора данных из Интернета или с камер видеонаблюдения с целью извлечения изображений лиц для создания или расширения баз данных, для слежения за людьми и распознавания эмоций на рабочем месте и в учебных заведениях. Однако здесь могут быть и исключения, так как иногда необходимо следить за человеком по медицинским показаниям или по соображениям безопасности (например, для контроля уровня усталости пилота). Регламент содержит и множество обязывающих предписаний. Так, например, он устанавливает, что поставщики систем ИИ с высоким уровнем риска будут должны самостоятельно оценивать и контролировать риски для здоровья, безопасности и основных прав. При этом необходимо обеспечить использование высококачественных данных для обучения алгоритмов и исключить предвзятость. Они же должны предоставлять техническую документацию и предоставлять точную и полную информацию организациям, осуществляющим внедрение (например, государственным органам, закупающим систему).

Регламент не регулирует защиту данных, деятельность онлайнплатформ и модерацию контента.. Регламент является дополнительным к законодательству о защите прав потребителей, в области защиты данных, безопасности продукции, основных прав работников . Основной принцип регулирования, используемый Регламентом, может быть отнесен к императивному принципу. Однако вместо прямых запретов использован рискоориентированный подход, который позволяет несколько ретушировать императивные нормы, четко определяя запреты только в отношении определенной группы систем ИИ – высокорисковых и неприемлемых.

Регламент ЕС об искусственном интеллекте признает обучение самостоятельными обязанностями:

- Article 9 (система управления рисками) предписывает, что для систем высокого риска управление рисками должно осуществляться на протяжении всего жизненного цикла. Это легитимирует ограничения на этапе обучения как часть общей стратегии;
- Article 10 (данные и управление данными) закрепляет требования к обучающим датасетам: они должны быть релевантными, репрезентативными, не содержать ошибок. Устанавливаются требования к практикам управления данными – прямое признание «датасет-уровня» регулирования;
- Article 53 (обязанности для провайдеров моделей общего назначения) обязывает разрабатывать политику соблюдения авторского права и публиковать подробный дайджест об использованном для обучения контенте;
- Article 113 (поэтапность применения) закрепляет отсрочку вступления в силу отдельных положений, что служит аргументом в пользу внедрения риск-ориентированного подхода с этапностью, а не тотальных запретов.

Из сферы применения Регламента исключены системы искусственного интеллекта, размещенные на рынке, введённые в эксплуатацию или используемые исключительно в военных целях, обороне или национальной безопасности, а также те системы ИИ, которые используются в личных целях, находятся в стадии разработки и ряд др. Все системы ИИ распределены должны быть по группам риска. Из них к группе минимального риска относятся: большинство систем ИИ, например, видеоигры. Разработчики и распространители таких систем ИИ не обязаны следовать требованиям Регламента. Однако это не исключает возможности разработчиков добровольно принимать свои правила и ограничения [15].

В США и Канаде практика варьируется в зависимости от юрисдикции полагаясь на существующие нормы в области защиты персональных данных и судебный прецедент. Использование чувственных данных, включая потенциально запрещенный контент, для тренировки алгоритмов обычно происходит в рамках закрытых программ с ограниченным доступом и внутренними протоколами этической проверки. В Канаде в 2022 году был принят Закон о внедрении Директивы о порядке использования ИИ в государственном секторе, обязывающий федеральные учреждения проводить оценку воздействия алгоритмов перед организацией автоматизированных систем, влияющих на граждан. Процедуры ее проведения остаются в ведении специализированных агентств и регулируются их внутренними директивами, общими нормами уголовного процессуального права.

Не смотря на активное внедрение обучения ИИ для государственных нужд, общей тенденцией в демократических правопорядках является движение к признанию необходимости специальных процедур, повышенной прозрачности и независимого аудита при использовании чувственных данных для обучения ИИ. Сам процесс создания ИИ для публичных нужд становится зоной правового напряжения, где сталкиваются императив безопасности и незыблемость частных прав.

Прогрессирующая сложность технологий неизбежно выдвигает на передний план вопрос о прозрачности алгоритмических решений, что, в свою очередь, актуализирует задачу поиска адекватных форм и границ правового воздействия.

Анализ российского и зарубежного права позволяет сформулировать систему принципов. На которых должно строиться правовое регулирование обучения ИИ в публичных целях:

- приоритет прав и свобод человека и гражданина – публичный интерес не может автоматически преодолевать частную автономию;
- принцип пропорциональности – тест на соразмерность вмешательства должен стать обязательным при решении вопроса об обучении ИИ на данных, затрагивающих права граждан. Если цель может быть достигнута с использованием обезличенных данных, использование реальных персональных данных недопустимо;
- принцип технологической нейтральности – требования должны формулироваться через описание стадий жизненного цикла и свойств информации, а не через привязку к конкретным технологиям;
- принцип прозрачности и подотчетности – разработчик и государственный заказчик должны быть способны продемонстрировать, на каких данных обучалась модель, какие меры фильтрации применялись, как исключалась предвзятость. Это предполагает введение обязательной документации датасета и процедур обучения.

Центральными вызовами для правовой доктрины, как мы полагаем, являются регулирование массивов данных и методов машинного обучения. Перед российским законодателем, по признанию экспертов, стоит нетривиальная задача: выработать такие правовые механизмы, которые бы гармонично сочетали защиту данных и стимулы для инноваций, не воздвигая при этом неоправданных барьеров. По нашему убеждению, искомый баланс достижим лишь при условии, что в основу нормативной рамки будут заложены принципы гибкости и адаптивности, позволяющие ей эволюционировать вместе с технологиями и предотвращать регуляторный застой. Одним из самых центральных вопросов является проблема ответственности, известная как «черный ящик». Ситуация усложняется, когда решение, повлекшее за собой ущерб для гражданина или общества, принимается автономной системы ИИ. Возникает вопрос: кто несет ответственность за это решение? Это может быть разработчик системы, оператор, который использует данную технологию, или государственный орган, который регулирует ее использование. На сегодняшний день существующие правовые конструкции плохо применимы к неличным субъектам, таким как искусственный интеллект. Сложность современных алгоритмов, которые основаны на методах глубокого обучения делают невозможным объяснение логики, по которой было принято то или иное решение. Если граждане не могут понять, почему было принято решение, которое затрагивает их права, это может подорвать доверие к правовой системе

Исследования В.Б. Наумова, С.А. Чеховской, А.Ю. Брагинца, А.В. Майорова определяют фундаментальные критерии, которым должны соответствовать многофункциональные системы искусственного интеллекта. Ключевым требованием выступает возможность перманентного контроля их функционирования, обеспечивающего прозрачность, интерпретируемость процесса принятия решений, возможность независимой верификации.

Институционализации, наряду с прочим, подлежат механизмы юридической ответственности за действия, совершаемые автономными системами, что является императивом для обеспечения высокой эффективности и предсказуемости правоприменения. Помимо неукоснительного соблюдения формально-правовых предписаний, особую значимость приобретает проблема имплементации этических императивов непосредственно в архитектуру систем. Фундаментальными требованиями должны стать обеспечение возможности оперативного корректирующего вмешательства человека (human-in-the-loop), а также создание многоуровневых гарантий отказоустойчивости (робастности) и информационной безопасности функционала подобных систем [13].

Умаление человеческого контроля, обусловленное передачей функций искусственному интеллекту, детерминирует необходимость ужесточения процедур верификации и мониторинга технологий. Нормативным воплощением этого подхода в России стал ФЗ «Об экспериментальных правовых режимах...» [1] — акт, не только определяющий цели и принципы, но и детально регламентирующий состав участников и порядок взаимодействия в рамках правовых «песочниц» для апробации инноваций. Концептуально цифровая трансформация представляет собой совокупность мер по модернизации госаппарата. Данный процесс имманентно предполагает, во-первых, ревизию нормативного массива на предмет выявления пробелов и устаревших норм; во-вторых, разработку нового регуляторного контура (защита данных, кибербезопасность); в-третьих, что является залогом успеха, — учреждение постоянно действующих экспертных площадок для институционализации взаимодействия государства, технологического сектора и науки.

Подобное взаимодействие создаёт условия для взаимного обмена профессиональным опы-

том и распространения наиболее эффективных методик работы. Это формирует фундамент для целостного и системного внедрения цифровых инструментов.

Разрешение проблемы так называемого «чёрного ящика», возникающей при использовании систем искусственного интеллекта в публичном пространстве, предполагает их осмысленное и строго регламентированное правом включение в общественные процессы. Классическая аргументация Герберта Саймона, указывающая на неизбежные когнитивные и мотивационные искажения в человеческих суждениях, подкрепляет тезис о том, что грамотно настроенный и надлежащим образом контролируемый искусственный интеллект способен выступать в роли более объективного вспомогательного механизма. Для того чтобы имманентно присущий технологиям ИИ потенциал был в полной мере реализован, непреложным императивом становится вовлечение законодателей и теоретиков права в процесс их создания еще на стадии пролегоменов — этапах концептуального проектирования и архитектурного моделирования. Именно на юридическое сообщество возлагается миссия по превентивному встраиванию в саму «ткань» алгоритма таких свойств, как транспарентность, логическая обоснованность, верифицируемость и интерпретируемость. Причем эти качества должны быть обеспечены на двух уровнях: как для узкопрофильных экспертов, способных провести технический аудит системы, так и для рядовых граждан, чьи права и законные интересы могут быть затронуты, и которые имеют право на получение мотивированного и понятного объяснения принятого машиной решения.

Проблема «черного ящика» из технической преобразуется в центральную задачу публичной власти, чье решение требует синергии нормотворчества, технологической стандартизации и формирования новой этико-правовой парадигмы. Вследствие этого требование понятности вычислительных процедур должно эволюционировать из технической спецификации в самостоятельное правовое начало. Нормативное закрепление этого начала — непременное условие законности, подотчетности власти и защиты прав граждан при автоматизации решений, примером чему служит Общий регламент по защите данных ЕС (GDPR), уже закрепивший право на объяснение решений, принятых машиной.

Следовательно, понятность вычислительных процедур недопустимо рассматривать как необязательную или второстепенную черту; она представляет собой нормативно значимую, системообразующую правовую категорию, требующую прямого и недвусмысленного законодательного определения. Принцип должен быть утвержден как всеобщее основание деятельности органов власти в цифровую эпоху, обеспечивающее правомерность (соответствие букве и духу закона), предсказуемость (возможность для человека предвидеть исход дела), проверяемость (возможность независимой оценки) и воспроизводимость (возможность повторного анализа) решений, выработанных с применением систем искусственного разума.

Практическое воплощение этого подхода в России просматривается в положениях Национальной стратегии развития искусственного интеллекта на период до 2030 года, где прямо указана необходимость «обеспечения понятности работы искусственного интеллекта и доверия к нему». Более того, в рамках опытных правовых режимов (так называемых «правовых песочниц»), вводимых Федеральным законом № 258-ФЗ, апробируются механизмы контроля за применением подобных технологий в таких чувствительных сферах, как, например, беспилотный транспорт или здравоохранение. Отсутствие же чётко установленного и юридически обеспеченного начала понятности неизбежно ставит под сомнение правовую чистоту отдельных решений и саму правомерность действий «цифрового государства», подрывая основы правопорядка и доверие общества к власти.

Подтверждением нарастающего осознания данной проблемы на высшем государственном уровне служит доктринальный анализ действующих документов стратегического планирования, таких как Национальная стратегия развития искусственного интеллекта. В них уже прямо артикулируется потребность в создании и институционализации конкретных правовых механизмов. Эти механизмы призваны обеспечить формальную (техническую) открытость кода, содержательную, то есть смысловую объяснимость логики функционирования систем ИИ, что является необходимым условием их интеграции в чувствительные сферы общественной жизни. Также отмечается важность внедрения стандартов интерпретируемости алгоритмов, формирования моделей юридической ответственности, защиты публичных интересов потенциальных рисков, порождаемых неконтролируемой алгоритмической непрозрачностью. Разработка систем, способных генерировать понятные человеку обоснования своих выводов, создаёт техническую базу последующей правовой подотчётности. Институционализация прозрачности именно как правового принципа становится, таким образом, необходимым условием легитимной интеграции в социальную среду технологий искусственного интеллекта. Формирование соответствующих регуляторных стандартов невозможно без прямого участия юридического сообщества, законо-

дателей, правоприменителей. Такие стандарты призваны гарантировать защиту прав личности, заложить основы общественного доверия к цифровым институтам государства, обеспечить устойчивое, отвечающее общественным ожиданиям их функционирование.

Поиск оптимального баланса между возможным преимуществами и рисками, связанными с внедрением искусственного интеллекта, требует не только технологических инноваций, но и правового и этического осмысления. Это значит, что необходимо разработать всестороннюю юридическую концепцию, которая будет регулировать использование ИИ в публичной сфере. Эта концепция должна включать несколько ключевых элементов, которые помогут создать безопасную и этическую среду для применения технологий. Во-первых, важно установить законодательные принципы, согласно которым ИИ будет находиться под контролем человека, что обеспечить возможность вмешательства в случае необходимости. Во-вторых, необходимо ввести презумпцию объяснимости и прозрачности алгоритмов, которые используются государственными органами. Это требует предоставления гражданам доступа к информации о том, как работают алгоритмы, какие данные используются и на каких основаниях принимаются те или иные решения. Прозрачность является ключевым аспектом, который обеспечивает доверии общества к технологиям. Третьим важным элементом является создание четких механизмов распределения ответственности за возможный ущерб. Который может быть причинен системами ИИ. Важно определить, кто будет нести ответственность в случае ошибок или злоупотреблений, связанных с использованием искусственного интеллекта, чтобы пострадавшие могли получить соответствующую компенсацию и защиту своих прав. Также следует ввести обязательные независимые аудиты, которые будут проверять системы ИИ на предмет дискриминации и соблюдения прав человека. Это поможет выявить и устранить потенциальные предвзятости в алгоритмах, которые могут негативно сказаться на определенных группах населения. Также необходимо обеспечить возможность публичного обсуждения и парламентского утверждения целей и ключевых параметров систем ИИ, которые будут внедряться в публичной сфере, что будет способствовать повышению уровня ответственности и подотчетности. Без наличия такой правовой рамки, создание ИИ, который должен служить публичному интересу, может трансформироваться в инструмент, защищающий интересы власти от общества, а не наоборот.

Регулирование искусственного интеллекта, непосредственно применяемого в публичной сфере, предполагает создание комплексной юридической и этической концепции. Эта концепция призвана обеспечить безопасное, подконтрольное и ответственное использование данных технологий. Основополагающим принципом выступает законодательное закрепление контроля человека над системами ИИ, гарантирующее возможность прямого вмешательства в критических ситуациях. Именно усиливать существующие демократические институты, служить дополнением правовым гарантиям (то есть не подменять) — такова единственно допустимая роль искусственного интеллекта. В конечном итоге, создание безопасного и этичного ИИ в публичной сфере является ообщей ответственностью правоведов, инженеров-разработчиков, политиков, всего общества становится построение безопасной, этичной системы.

Список литературы:

[1] Об экспериментальных правовых режимах в сфере цифровых инноваций в РФ : Федер. закон от 31.07.2020 г. N 258-ФЗ // СЗ РФ. – 2020. – N 31 (ч. I). – Ст. 5017.

[2] О внесении изменений в Указ Президента РФ от 10.10.2019 г. N 490 «О развитии искусственного интеллекта в РФ»... : Указ Президента РФ от 15.02.2024 г. N 124 // СЗ РФ. – 2024. – N 8. – Ст. 1102.

[3] О развитии искусственного интеллекта в РФ (вместе с «Национальной стратегией...») : Указ Президента РФ от 10.10.2019 г. N 490 // СЗ РФ. – 2019. – N 41. – Ст. 5700.

[4] Алиев И. М. Влияние технологий ИИ на рынок труда в России // Журн. правовых и экон. исследований. – 2019. – № 4. – С. 7-12.

[5] Балашов А. А. Концептуальные основы использования ИИ... // Цифровое моделирование экономики. – 2024. – № 1. – С. 74-79.

[6] Бодров Е. Н. К вопросу применения компетентностного подхода... // Безопасность. Управление. Искусственный интеллект. – 2022. – Т. 3, № 3. – С. 45-50.

[7] Выявление направлений использования ИИ в правоприменительной деятельности / А. Р. Мустафин [и др.] // Юрид. наука. – 2023. – № 5. – С. 17-20.

[8] Германов Н. С. Концепция ответственного ИИ... // Digital Diagnostics. – 2023. – Т. 4, № S1. – С. 27-29.

[9] Катанандов С. Л. Технологическое развитие современных государств: ИИ в госуправлении // Гос. и муницип. управление. Учен. записки. – 2023. – № 1. – С. 174-182.

[10] Кирилова Е.А., Зульфугарзаде Т.Э. К вопросу о правосубъектности ИИ // Имущественные отношения в РФ. – 2024. – N 4. – С. 76-80.

[11] Лащенов М. С. Эмоциональный интеллект в управленческой деятельности и технологиях ИИ // Цифровая социология. – 2024. – Т. 7, № 4. – С. 44-52.

[12] Мартынова Ю. А. Развитие и использование методов ИИ в различных сферах... // Экономика и управление: проблемы, решения. – 2024. – Т. 9, № 6. – С. 187-193.

[13] Правовые аспекты использования искусственного интеллекта... : доклад НИУ ВШЭ / В.Б. Наумов [и др.]. – М. : ИД ВШЭ, 2021. – 42 с.

[14] Овсянникова М. В. Понятие «искусственный интеллект» и его использование... // Проблемы науч. мысли. – 2024. – Т. 12, № 4. – С. 158-160.

[15] Овчинников А. И. Регламент Евросоюза об искусственном интеллекте: эксперимент в сфере правового регулирования цифровых технологий // Юридическая техника. – 2025. – № 19: Материалы XXVI Международного междисциплинарного научно-практического форума «Юртехнетика» на тему «Эксперимент в правовом регулировании (доктрина, практика, техника)» (Нижний Новгород, 25–28 сентября 2024 года) / гл. ред. проф. В. М. Баранов. – Нижний Новгород: ЮНИКОПИ, 2025. – С. 207-209.

[16] Шафалович А. А. Внедрение ИИ в нормотворческую деятельность... // Право.by. – 2023. – № 5. – С. 13-19.

[17] Шиткина И.С., Бирюков Д.О. Искусственный интеллект: правовые аспекты // Право и экономика. – 2023. – N 11.

[18] Человек vs робот: надо ли выбирать? – URL: https://spb.plus.rbc.ru/news/66729c767a8aa98aab28d0ad (дата обращения: 26.05.2024).

References:

[1] On experimental legal regimes in the field of digital innovation in the Russian Federation: Feder. Law of 31.07.2020 No. 258-FZ//CZ of the Russian Federation. – 2020. - N 31 (part I). - Art. 5017.

[2] On amendments to the Decree of the President of the Russian Federation of 10.10.2019 N 490 “On the development of artificial intelligence in the Russian Federation...”: Decree of the President of the Russian Federation of 15.02.2024 N 124//SZ RF. – 2024. – N 8. - Art. 1102.

[3] On the development of artificial intelligence in the Russian Federation (together with the “National Strategy...”): Decree of the President of the Russian Federation of 10.10.2019 N 490//NW RF. – 2019. – N 41. - Art. 5700.

[4] Aliyev I. M. The influence of AI technologies on the labor market in Russia//Journal. legal and economy. studies. – 2019. – № 4. - S. 7-12.

[5] Balashov A. A. Conceptual foundations of the use of AI ...//Digital modeling of the economy. – 2024. – № 1. - S. 74-79.

[6] Bodrov E.N. To the issue of applying a competency approach ...//Safety. Management. Artificial intelligence. – 2022. - T. 3, NO. 3. - S. 45-50.

[7] Identification of areas of AI use in law enforcement/A. R. Mustafin [et al.]//Yurid. science. – 2023. – № 5. - S. 17-20.

[8] Germanov N. S. The concept of responsible AI ...//Digital Diagnostics. – 2023. - VOL. 4, NO. S1. - S. 27-29.

[9] Katanandov S. L. Technological development of modern states: AI in public administration// State. and munitif. management. Learned. notes. – 2023. – № 1. - S. 174-182.

[10] E.A. Kirilova, T.E. Zulfugarzade. On the issue of AI legal personality//Property relations in the Russian Federation. – 2024. – N 4. - S. 76-80.

[11] Lashchenov M.S. Emotional intelligence in management activities and AI technologies//Digital sociology. – 2024. - T. 7, NO. 4. - S. 44-52.

[12] Martynova Yu. A. Development and use of AI methods in various fields ...//Economics and management: problems, solutions. – 2024. - T. 9, NO. 6. - S. 187-193.

[13] Legal aspects of the use of artificial intelligence...: HSE report/V.B. Naumov [et al.]. - M.: HSE PUBLISHING HOUSE, 2021. - 42 s.

[14] Ovsyannikova M.V. The concept of “artificial intelligence” and its use ...//Problems scientific. thoughts. – 2024. - T. 12, NO. 4. - S. 158-160.

[15] Ovchinnikov A.I. Regulations of the European Union on artificial intelligence: an experiment in the field of legal regulation of digital technologies// Legal technology. – 2025. - No. 19: Materials of the XXVI International Interdisciplinary Scientific and Practical Forum “Yurtekhnetics” on the topic “Experiment in legal regulation (doctrine, practice, technique)” (Nizhny Novgorod, September 25-28, 2024)/ Ch. ed. Prof. V. M. Baranov. - Nizhny Novgorod: UNICOPI, 2025. - S. 207-209.

[16] Shafalovich A. A. Introduction of AI into normative activities ...// Право.by. – 2023. – № 5. - S. 13-19.

[17] Shitkina I.S., Biryukov D.O. Artificial intelligence: legal aspects//Law and economics. – 2023. – N 11.

[18] Man vs robot: do I have to choose? - URL: https://spb.plus.rbc.ru/news/66729c767a8aa98aab28d0ad (accessed: 26.05.2024).