Дата поступления рукописи в редакцию: 27.03.2026 г.
Дата принятия рукописи в печать: 30.04.2026 г.
DOI: 10.24412/2076-1503-2026-4-767-774

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

ОБУЧЕНИЕ ИИ НА ИНФОРМАЦИИ ОГРАНИЧЕННОГО ДОСТУПА И ЗАПРЕТ НА ИСПОЛЬЗОВАНИЕ ЗАПРЕЩЁННОЙ ИНФОРМАЦИИ: ОСНОВАНИЯ ОБРАБОТКИ, ГАРАНТИИ, ПУБЛИЧНО ПРАВОВАЯ ОТВЕТСТВЕННОСТЬ СУБЪЕКТОВ

Аннотация. *В статье исследуются правовые основания и пределы использования информации ограниченного доступа при обучении систем искусственного интеллекта, а также публично-правовые механизмы недопущения включения в датасеты запрещённой информации. Обосновывается, что обучение моделей ИИ следует квалифицировать как форму обработки персональных данных, вследствие чего на соответствующие процессы распространяются общие принципы законности, целевого характера, минимизации и безопасности обработки. Особое внимание уделяется сопоставлению российского регулирования с подходами Европейского союза, включая положения GDPR и AI Act, устанавливающими требования к прозрачности, качеству обучающих данных, управлению рисками и документированию источников контента. Выявляется наличие пробела в российском законодательстве, связанного с отсутствием специального правового основания для обработки персональных данных в целях обучения ИИ без получения согласия каждого субъекта. В качестве решения предлагается закрепление в законодательстве самостоятельного основания такой обработки при условии соблюдения минимального стандарта гарантий, включающего обезличивание и псевдонимизацию данных, документирование происхождения и состава датасетов, оценку рисков, запрет реидентификации и обязательную фильтрацию запрещённого контента. Делается вывод о необходимости перехода к модели ответственности, ориентированной прежде всего на соблюдение оператором процедурных обязанностей.*

Ключевые слова: *искусственный интеллект; персональные данные; обработка персональных данных; информация ограниченного доступа; датасет; обучение искусственного интеллекта; запрещённая информация; правовые основания обработки; обезличивание данных; EU AI Act; GDPR; публично-правовая ответственность; фильтрация контента; управление рисками.*

CHUMACHENKO Vasily Andreevich,
3 postgraduate course St. Petersburg State University

TRAINING ON RESTRICTED INFORMATION AND PROHIBITION OF USE OF PROHIBITED INFORMATION: GROUNDS FOR PROCESSING, GUARANTEES, PUBLIC LEGAL LIABILITY OF SUBJECTS

Annotation. *The article examines the legal grounds and limits of the use of restricted access information in the training of artificial intelligence systems, as well as public law mechanisms to prevent the inclusion of prohibited information in datasets. It is proved that the training of AI models should be qualified as a form of personal data processing, as a result of which the general principles of legality, purposefulness, minimization and security of processing apply to the relevant processes. Particular attention is paid to comparing Russian regulation with the approaches of the European Union, including the provisions of the GDPR and the AI Act, which establish requirements for transparency, quality of training data, risk management and documentation of content sources. There is a*

gap in Russian legislation related to the lack of a special legal basis for processing personal data for the purpose of AI training without obtaining the consent of each subject. As a solution, it is proposed to consolidate in legislation an independent basis for such processing, subject to compliance with the minimum standard of guarantees, including depersonalization and pseudonymization of data, documentation of the origin and composition of datasets, risk assessment, prohibition of re-identification and mandatory filtering of prohibited content. It is concluded that it is necessary to switch to a responsibility model focused primarily on the operator's compliance with procedural duties.

Key words*: artificial intelligence; personal data; processing of personal data; restricted access information; dataset; artificial intelligence training; prohibited information; legal grounds for processing; depersonalization of data; EU AI Act; GDPR; public liability; content filtering; risk management.*

Развитие технологий искусственного интеллекта, особенно методов машинного обучения, используемых при создании генеративной модели, ставит перед правовой системой принципиально новые задачи, связанные с оборотом информации [6].

Подход машинного обучения (Machine Learning, ML) в отличие от систем на основе правил не использует заранее определённые инструкции. Система самостоятельно выявляет закономерности и зависимости из данных. Обучаясь на примерах, модель способна делать прогнозы и принимать решения без прямого вмешательства человека. Таким образом, машинное обучение обеспечивает гибкость и адаптивность, позволяя создавать интеллектуальные системы, способные решать сложные задачи, которые невозможно полностью описать с помощью заранее заданных правил. Этот подход является фундаментом современных приложений ИИ, от рекомендательных систем и автономных транспортных средств до прогнозной аналитики [8].

Для эффективного обучения AI-систем требуется огромный объём качественной информации. Если ранее законодательство и правоприменительная практика были сосредоточены главным образом на регулировании распространения готовой информации, то сегодня важное значение отведено стадии формирования наборов данных – датасетов. Датасет как технический агрегатор огромного массива сведений, в том числе персональных данных и информации ограниченного доступа, превращается в объект, требующий публично-правового контроля. Качество его наполнения и процедуры обработки напрямую предопределяют потенциальные риски для прав граждан, безопасности государства и общественной нравственности. Глубокие нейронные сети нередко рассматриваются как «чёрные ящики»: они показывают отличные результаты, но их внутренние решения трудно объяснить. В связи с этим перед юридической наукой и законодателем встает ключевой вопрос: каким образом легализовать или обосновать запрет использования персональных данных и иной охраняемой информации для обучения искусственного интеллекта, а также с помощью каких правовых механизмов гарантировать недопущение в датасетах запрещенного контента, обеспечив при этом проверяемость соответствующих процедур со стороны надзорных органов [4].

Успех всех нейронных сетей напрямую зависит от качества данных, на которых они обучаются. Проблемы, связанные со сбором и обработкой данных, существенно снижают точность и эффективность моделей ИИ. Во всех упомянутых областях данные, используемые для обучения моделей ИИ, играют ключевую роль в определении их эффективности и точности. Поэтому создание качественных датасетов играет ключевую роль в разработке успешных моделей искусственного интеллекта. Проблема сбора данных заключается в нехватке информации и ограниченном разнообразии источников. Это приводит к снижению репрезентативности датасета. Также следует учитывать, что для такого объѐмного сбора данных потребуется значительное количество специалистов. После сбора информации необходима предобработка данных. Неправильные или пропущенные данные могут сильно искажать результаты обучения. Проблемы в большинстве случаев возникают при очистке и нормализации данных, а также разные форматы данных могут затруднить их объединение и использование, требуя дополнительных шагов по стандартизации. Одной из ключевых проблем является репрезентативность и баланс данных. При обучении нейронной сети необходимо соблюдать классовый баланс, иначе модель может стать предвзятой, а недостаток разнообразия в данных, в свою очередь, ограничивает способность модели обобщать информацию и адаптироваться к новым ситуациям. Для качественного тестирования нейросеть должна быть обучена на разнообразных данных, которые отражают реальные условия функционирования системы [1].

Одна из главных проблем при использовании данных для обучения модели – это вопрос, как правомерно использовать эти данные с соблюдением всех законов и актов. Обучение модели искусственного интеллекта, включая процессы сбора, систематизации, накопления и использова-

ния данных для выявления закономерностей и формирования математических зависимостей, по своей природе представляет собой одну из форм обработки персональных данных. Согласно статье 3 Федерального закона от 27 июля 2006 №152-ФЗ «О персональных данных», под обработкой понимается любое действие или совокупность действий, совершаемых с использованием средств автоматизации, включая запись, систематизацию, накопление, хранение, уточнение, извлечение и использование. Все перечисленные операции неизбежно сопровождают процесс подготовки и тренировки нейросетей. Следовательно на операторов, осуществляющих такое обучение, распространяются общие принципы обработки персональных данных, закрепленные в статье 5 Закона о персональных данных. Среди них особое значение приобретают принципы обработки только в заранее определенных и законных целях, а также принцип минимизации, согласно которому обрабатываемые данные не должны быть избыточными по отношению к заявленным целям.

Интерес представляет опыт правового регулирования этих вопросов в Европейском союзе. В Европейском союзе регулирование обработки персональных данных базируется на Общем регламенте по защите данных (GDPR), который гарантирует субъектам целый ряд прав, включая право не подвергаться решениям, полностью принимаемых алгоритмом и оказывающих существенное влияние на их положение [7].

Общий регламент по защите данных (GDPR) применяется ко всем организациям, которые обрабатывают данные резидентов ЕС, включая иностранных разработчиков ИИ. Он требует высокой прозрачности обработки: субъект должен быть информирован об операторе, целях и правовом основании обработки, категориях данных, сроках их хранения, возможных получателях, включая зарубежных, а также о факте использования автоматизированных решений. GDPR закрепляет широкий набор прав граждан. Среди них право доступа к данным, их исправления и удаления, ограничение обработки, переносимость и возможность возражения. Особое значение имеет статья 22, устанавливающая запрет на полностью автоматизированные решения, которые существенно влияют на положение человека, если отсутствует его явное согласие или иное законное основание. Даже в исключительных случаях гражданин сохраняет право требовать пересмотра решения с участием человека. Такой подход гарантирует, что алгоритмы не действуют автономно и остаются под контролем [3].

Важнейшим событием стало вступление в силу Регламента ЕС об искусственном интеллекте (AI Act). Регламент был опубликован 12 июля 2024года и вступил в силу 1 августа 2024 года, однако отдельные положения применяются поэтапно: запреты на системы с неприемлемым риском действуют с 2 февраля 2025 года, однако отдельные положения применяются поэтапно: запреты на системы с неприемлемым риском действуют с 2025 года, а требования к системам общего назначения (GPAI), включая обязательства по прозрачности обучающих данных, - с 2 августа 2025 года. Данный документ относится к области технического регулирования и сосредотачивается на безопасности и надежности ИИ-систем, не расширяя перечень прав субъектов данных, но устанавливая важные требования к стадии обучения.

С точки зрения обучения ИИ, ключевое значение имеет статья 10 AI Act, посвященная данным и управлению данными для систем высокого риска. Она устанавливает императивные требования к обучающим, валидационным и тестовым наборам данных: они должны быть релевантными, репрезентативными, максимально свободными от ошибок и полными с учетом предполагаемого назначения системы. Статья обязывает применять практики управления данными, включающие анализ возможных предвзятостей и принятие мер для их обнаружения и предотвращения. Особо оговаривается, что обработка специальных категорий персональных данных (например, о расовой принадлежности, политических взглядах, здоровье) допускается лишь в исключительных случаях, строго необходимых для обеспечения выявления и коррекции предвзятости, и при условии соблюдения усиленных мер защиты.

Для систем ИИ общего назначения (GPAI) статья 53 (1)(d) AI Act вводит требование о публикации подробного резюме использованного для обучения контента. Во исполнение этой нормы Европейская комиссия в июле 2025 года утвердила обязательный шаблон такого резюме, который вступил в силу для новых моделей со 2 августа 2025 года. Шаблон требует раскрытия общей информации об обучающих данных, перечня источников (включая общедоступные датасеты, данные из интернета с указанием основных доменов, лицензированные частные данные) и информации о мерах по удалению нелегального контента и соблюдению прав интеллектуальной собственности (в частности, механизмов opt-out, предусмотренных Директивой об авторском праве) [9].

Важно подчеркнуть, что AI Act не отменяет, а дополняет требования GDPR. Европейский совет по защите данных в своем Заключении 28/2024 разъяснил, что использование персональных данных для обучения моделей ИИ подпадает под действие GDPR. В частности, в Заключении ука-

зывается, что законный интерес не может быть универсальным основанием для обработки и требует тщательной проверки по трехступенчатому тесту. Модель ИИ, обученная на персональных данных, не может автоматически считаться анонимной – анонимность должна быть доказана с учетом риска извлечения данных из модели.

Параллельно формируется более широкая международная база: обновленная Конвенция Совета Европы, Принципы ОЭСР по ответственному ИИ и Рекомендации ЮНЕСКО по этике ИИ последовательно подчеркивают важность защиты прав человека и приватности при разработке и использовании алгоритмических технологий.

GDPP применяется ко всем организациям, которые обрабатывают данные резидентов ЕС, включая иностранных разработчиков ИИ. Он требует высокой прозрачности обработки: субъект должен быть информирован об операторе, целях и правовом основании обработки, категориях данных, сроках их хранения, возможных получателях, включая зарубежных, а также о факте использования автоматизированных решений. GDPR закрепляет широкий набор прав граждан. Среди них право доступа к данным, их исправления и удаления, ограничение обработки, переносимость и возможность возражения. Особое значение имеет статья 22, устанавливающая запрет на полностью автоматизированные решения, которые существенно влияют на положение человека, если отсутствует его явное согласие или иное законное основание. Даже в исключительных случаях гражданин сохраняет право требовать пересмотра решения с участием человека. Такой подход формирует гарантию того, что алгоритмы не действуют автономно и остаются под контролем.

В свою очередь российский законодатель в действующей редакции статьи 6 Закона о персональных данных, устанавливающей перечень оснований для обработки персональных данных, не сформулировал специального положения, которое позволило бы однозначно квалифицировать обучение искусственного интеллекта как правомерную деятельность без получения согласия каждого субъекта. Такие основания, как исполнение договора, достижение законных интересов или обработка общедоступных данных, не покрывают всей специфики тренировочных процессов, особенно когда данные собираются путем автоматизированного скрапинга из открытых источников. Получение же информационного и конкретного согласия от миллиона пользователей, чьи данные потенциально могут войти в датасет, технически невозможно и экономически нецелесообразно. Данный пробел создает правовую неопределенность, которая усугубляется общим трендом на ужесточение требований к локализации и защите данных, нашедшим отражение, в частности, в Федеральном законе от 14 июля 2022 г. № 237-ФЗ “О внесении изменений в отдельные законодательные акты Российской Федерации”, установившем запрет на сбор персональных данных граждан Российской Федерации с использованием зарубежных баз данных. Это свидетельствует о том, что законодатель рассматривает любые операции с данными, включая обучение алгоритмов, как часть национального правового поля, требующую четкой регламентации.

Выход из сложившейся ситуации видится в дополнении статьи 6 Закона № 152-ФЗ новым самостоятельным основанием, специально предназначенным для целей обучения искусственного интеллекта. Представляется целесообразным ввести в закон часть 2.1. следующего содержания: «Обработка персональных данных для целей обучения (тренировки) моделей искусственного интеллекта, включая создание и тестирование наборов данных (датасетов), допускается без получения согласия субъекта персональных данных при условии обеспечения оператором минимального стандарта гарантий и невозможности использования данных для иных целей». Подобная конструкция позволяет отказаться от требования индивидуального согласия, одновременно возлагая на оператора публично-правовые обязанности по соблюдению процедурных гарантий, призванных обеспечивать защиту прав субъектов.

Содержание минимального стандарта гарантий должно быть раскрыто либо непосредственно в законе, либо в подзаконных актах, принимаемых уполномоченным органом. При его формировании целесообразно учитывать как зарубежный опыт, в частности положения статьи 35 Общего регламента о защите данных (GDPR), предписывающей проведение оценки воздействия на защиту данных (DPIA), так и разъяснения Европейского комитета по защите данных о том, что обучение моделей является составной частью обработки, требующей особых мер предосторожности. Исходя их этого, в минимальный стандарт гарантий должны быть включены следующие обязательные элементы. Во-первых, требование обезличивания или псевдонимизации данных на этапе включения в датасет, что снижает риск прямой идентификации субъектов. Во-вторых, принцип технологической необходимости: оператор обязан обосновать, что использование именно персональных данных является необходимым для достижения заявленных целей обучения, например, для задач распознавания голоса или поведенческого анализа. В-третьих, введение запрета на реидентификацию, подкрепленного техническими и организационными мерами, включая защиту от атак извлечения. В-четвертых, обя-

зательное документирование состава и происхождения датасета (введение так называемых Data Cards или Model Cards), что позволит в любой момент продемонстрировать надзорным органам, какие данные и на каком основании были использованы. В-пятых, проведение внутренней оценки рисков для прав субъектов до начала обучения, аналогичной DPIA.

Помимо защиты персональных данных критически важной задачей является предотвращение попадания в тренировочные наборы информации, оборот которой в Российской Федерации прямо запрещен или ограничен. Речь идет о материалах, включенных в федеральный список экстремистских материалов, ведение которого осуществляется Министерством юстиции на основании статьи 13 Федерального закона от 25 июля 2002 г. №114-ФЗ «О противодействии экстремистской деятельности», а также об информации, способной причинить вред здоровью и развитию детей, согласно Федеральному закону от 29 декабря 2010 г. №436-ФЗ. Основными критериями являются вовлечение несовершеннолетних в противоправную деятельность или навязывание деструктивных образцов и моделей поведения. Приказ Росздравнадзора от 29.06.2020 №5527 регулирует сферу незаконных продаж лекарств с помощью Интернета, что также считается вовлечением в противоправные деяния. По тематикам инструкций по созданию оружия и взрывчатки приказ запрещенная информация также направлена на вовлечение в противоправную деятельность неустановленного круга лиц на основании Приказа МВД России, ФСБ России, Росгвардии от 27.07.2022 №463/319/206. Перечень ФСЭМ (Федеральный Список Экстремистских материалов) регулируется удалением конкретных противоправных единиц контента, которые были признаны запрещенными по решению суда. Федеральный закон № 479-ФЗ «О внесении изменений в Кодекс Российской Федерации об административных правонарушениях» от 05.12.2022 г. вводит понятия запрещенной информации касательно тематики пропаганды ЛГБТ, смены пола и педофилии, что является борьбой с введением деструктивных явлений в понятие социальной нормы Постановление Правительства РФ от 02.11.2021 №1905 направлено на информацию, которая позволяет получить доступ к заблокированным ресурсам, т.е. получить доступ к противоправному контенту. Приказ Роскомнадзора от 27.02.2023 №25 «Об утверждении критериев…» выводит критерии запрещенной информации по тематикам детской порнографии, пронаркотического и суицидального контента, пропаганды нетрадиционных сексуальных отношений. Деструктивным контентом могут быть текста, фото, видеоматериалы, которые располагаются на интернет-страницах, социальных сетях, а также в СМИ. Также в приказе задержатся общая работа по материалам, содержащим запрещенную информацию. Приказ МВД России от 03.03.2023 № 114 более детально описывает запрещенную информацию по тематике пронаркотического контента. Согласно Приказу Роспотребнадзора от 27.02.2024 №79 в рамках критериев противоправной информации суицидального контента появляется тезис о том, какая информация не признается запрещенной: контент с признаками на нарушение обладает исторической, научной, художественной или иной ценностью для общества [5].

Поскольку обученная модель впоследствии генерировать подобный контент или воспроизводить его фрагменты, ответственность разработчика не может ограничиваться лишь реакцией на факты выдачи. Юридически значимым должно признаваться принятие оператором исчерпывающих мер по фильтрации датасетов. Такие меры можно разделить на два уровня: предварительная фильтрация, заключающаяся в отчистке исходных текстов от совпадений с хеш-суммами материалов из официальных реестров и использовании стоп-листов, и последующая контекстная фильтрация, предполагающая контекстная фильтрация, предполагающая тонкую настройку модели для обработки генерации запрещенного контента при отсутствии его прямых упоминаний в исходных данных. Доказательством добросовестности оператора должно служить наличие задокументированных процедур проверки датасета по официальным базам и классификаторам, посредством применения хеш-таблиц для идентификации и блокировки объекта, соответствующего признаками запрещенной информации.

Вопрос об ответственности за нарушения, допущенные при обучении искусственного интеллекта также требует переосмысления. Традиционный подход, при котором надзорные органы оценивают исключительно факт состоявшегося нарушения (например, обнаружение персональных данных в открытом датасете), не учитывает специфику технологических процессов, где абсолютное исключение ошибок невозможно в условиях максимальной предусмотрительности. Наиболее эффективной представляется модель, при которой объектом контроля становится не столько результат, сколько соблюдение процедурных обязанностей. Иными словами контролирующий орган должен проверять наличие и качество исполнения оператором минимальных стандарта гарантий: проводилась ли оценка рисков, велись ли логии, применялись ли меры обезличивания и фильтрации. Если оператор добросовестно внедрил все требуемые процедуры, но единичное

нарушение произошло в силу объективных и неустранимых технических причин, меры публично-правовой ответственности могут быть смягчены. Напротив, отсутствие задокументированных процедур и игнорирование требований к фильтрации должно квалифицироваться как грубое нарушение, в том числе, если обученная модель не породила вредоносных последствий. Такой подход требует внесения соответствующих изменений в Кодекс об административных правонарушениях – дополнения статьи 13.11 составом, предусматривающим ответственность за невыполнение оператором обязанности по документированию датасета и оценке рисков при обработке в целях обучения искусственного интеллекта.

Важно сказать, что сбор и хранение больших объемов конфиденциальных данных может создать основные цели для кибератак, что приведет к увеличению рисков безопасности. Хотя ИИ и МО значительно улучшили защиту безопасности, они также вносят новые уязвимости. Например, атаки состязательного ИИ включают создание входных данных, специально предназначенных для обмана систем ИИ, тем самым обходя меры безопасности, управляемые ИИ. Кроме того, если злоумышленники манипулируют данными обучения моделей ИИ, они могут вызывать предвзятость или ошибки, которые снижают эффективность моделей, делая их менее восприимчивыми или даже невосприимчивыми к реальным угрозам. Существует также риск снижения человеческого контроля при чрезмерной зависимости от ИИ и автоматизированных процессов, что впоследствии может привести к значительным пробелам в безопасности, особенно с новыми или сложными угрозами, которые не охвачены в наборах данных обучения. Для снижения этих рисков решающее значение имеют непрерывный мониторинг и регулярные обновления моделей ИИ для адаптации к новым и развивающимся угрозам. Внедрение гибридного подхода, объединяющего вычислительную мощность ИИ с человеческим опытом, может повысить надежность систем безопасности, гарантируя, что необычные или новые угрозы будут адекватно устранены. Более того, соблюдение этических практик разработки ИИ, включая поддержание прозрачности и приоритет безопасности, имеет важное значение для поддержания целостности и эффективности приложений ИИ в области безопасности, а также гарантируют, что развертывание технологий ИИ не ставят под угрозу целостность и безопасность [8].

Проведенный анализ позволяет говорить о том, что в признании обучения систем искусственного интеллекта на больших массивах данных, включающих персональные данные и иную охраняемую законом информацию, объективной технологической реальностью современного этапа цифрового развития. Персональные данные сегодня выступают стратегическим ресурсом государства и общества, однако их использование для целей развития технологий ИИ требует выработки единого системного правового подхода. Де-юре обучении ИИ признается формой обработки персональных данных и должно осуществляться в строгом соответствии с требованиями законодательства. В современной действительности невозможно и нецелесообразно установления презумптивного запрета на такое обучение. Запрет «по умолчанию» не только противоречит стратегическим задачам технологического развития, сформулированным в документах стратегического планирования, но и технически нереализуемым баз полной остановки развития цифровой экономики и сопряжен с риском утраты технологического суверенитета [10]. Как указывается в официальных источниках, целью государственной политики в сфере интеллектуальной собственности и цифровых инноваций является повышение конкурентоспособности экономики Российской Федерации, обеспечение роста ВВП, национальной безопасности и технологической независимости в важных для государства и общества сферах. Достижение этих целей невозможно без легализации и упорядочения процессов обучения ИИ. Принципиальным условием легализации соответствующих процессов выступает неукоснительное соблюдение операторами публично-правовых условий, образующих минимальный стандарт гарантий. Данный стандарт представляет собой систему императивных требований, обеспечивающих баланс между инновационным развитием и защитой конституционных прав граждан. При разработке новых методов охраны интеллектуальной собственности в цифровой среде принципиально важно обеспечить именно этот баланс. Минимальный стандарт гарантий включает следующие обязательные элементы, находящие свое нормативное обоснование в действующем и перспективном законодательстве. Во-первых, обязанность обезличивания персональных данных: Федеральным законом от 8 августа 2024 года №233-ФЗ «О внесении изменений в Федеральный закон «О персональных данных» и Федеральный закон «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» уже создан механизм формирования обе-

зличенных составов данных (датасетов) в государственной информационной системе, что направлено на обеспечение обмена обезличенными данными между бизнесом и государством, обучение моделей искусственного интеллекта, снижение объема обрабатываемых персональных данных и повышение их защищенности, при этом не допускается формирование составов данных из специальных категорий персональных данных (состояние здоровья, расовая и национальная принадлежность). Во-вторых, документирование происхождения и состава датасетов: введение обязанности ИТ-компаний раскрывать подробный «паспорт» каждого используемого датасета, включая наименование, формат. Объем, назначение, источник и дату создания, что корреспондирует с положениями статьи 18.1 Закона о персональных данных, обязывающей оператора иметь документы, определяющие политику в отношении обработки персональных данных. В-третьих, обязанность фильтрации запрещенного контента, необходимость которого подтверждается практикой Роскомнадзора, активно внедряющего с 2025 года технологии машинного обучения для поиска и блокировки запрещенной информации. Это требует от операторов применять аналогичные механизмы предварительной и последующей фильтрации, проверяя датасеты по официальным реестрам и классификаторам. В-четвертых, управление рисками искусственного интеллекта с целью выявления специфических для ИИ опасностей, сопоставления их с жизненным циклом ИИ и документирования элементов управления.

Список литературы:

[1] Бадаква, А. С. Создание наборов данных для искусственного интеллекта: ключевые вызовы и стратегии их преодоления / А. С. Бадаква, Б. В. Мартынов // Интеллектуальные ресурсы - региональному развитию. – 2024. – № 3. – С. 19-28.

[2] Брокман, Д. Что мы думаем о машинах, которые думают: Ведущие мировые ученые об искусственном интеллекте/Д. Брокман. М.: - Альпина нон-фикшн, 2017. - 552 с.

[3] Воробьева, А. А. Автоматизированный аудит информационной безопасности с использованием машинного обучения: вызовы для обеспечения безопасности ИИ / А. А. Воробьева // Журнал высоких гуманитарных технологий. – 2025. – № 3(10). – С. 29-39.

[4] Кожевникова, Я. И. Искусственный интеллект и машинное обучение: Объяснимый ИИ (XAI) и прозрачность алгоритмов / Я. И. Кожевникова, А. С. Белоусова // Аллея науки. – 2025. – Т. 1, № 4(103). – С. 824-827.

[5] Наумов, В. Б. Определение понятия запрещенной информации по тематикам Единого Реестра в рамках действующего законодательства РФ / В. Б. Наумов // Гуманитарный научный журнал. – 2025. – № 2-1. – С. 63-69.

[6] Сыдыкова, М. Б. Взаимосвязь между искусственным интеллектом, машинным обучением и глубоким обучением / М. Б. Сыдыкова, Г. К. Эсенаманова // Бюллетень науки и практики. – 2025. – Т. 11, № 12. – С. 474-480.

[7] Clark J., Demircan M., Kettas K. Europe: the EU AI Act's relationship with data protection law – key takeaways. – DLA Piper, Privacy Matters Blog, 25 апреля 2024 г.

[8] Технологические вызовы внедрения искусственного интеллекта и машинного обучения в безопасности: баланс между инновациями и ответственностью / Н. Ш. Козлова, В. А. Довгаль, А. К. Доргушаова, Р. С. Козлов // Вестник Адыгейского государственного университета. Серия: Естественно-математические и технические науки. – 2025. – № 1(356). – С. 55-69.

[9] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance) // https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng (дата обращения: 17.03.2026)

[10] Титков, Д. И. Резниченко, С.А. Революция в аудите информационной безопасности: искусственный интеллект /Д.И. Титков// Вестник науки, 2024. – Т. 1. № 4(73). – С. 431-438.

References:

[1] Badakva, A. S. Creation of data sets for artificial intelligence: key challenges and strategies for overcoming them/A. S. Badakva, B. V. Martynov// Intellectual resources - regional development. – 2024. – № 3. - S. 19-28.

[2] Brockman, D. What we think of machines that think: The world's leading scientists on artificial intelligence/D. Brockman. M.: - Alpina non-fiction, 2017. - 552 s.

[3] Vorobyova, A. A. Automated audit of information security using machine learning: challenges to ensure the security of AI/A. A. Vorobyov//Journal of High Humanitarian Technologies. – 2025. – № 3(10). - S. 29-39.

[4] Kozhevnikova, Ya. I. Artificial intelligence and machine learning: Explainable AI (XAI) and transparency of algorithms/Ya. I. Kozhevnikov, A. S. Belousov//Alley of Science. – 2025. - T. 1, NO. 4 (103). - S. 824-827.

[5] Naumov, V. B. Definition of the concept of prohibited information on the topics of the Unified Register within the framework of the current legislation of the Russian Federation/V. B. Naumov// Humanitarian Scientific Journal. – 2025. – № 2-1. - S. 63-69.

[6] Sydykova, M. B. The relationship between artificial intelligence, machine learning and deep learning/M. B. Sydykova, G. K. Esenamanova//Bulletin of Science and Practice. – 2025. - T. 11, NO. 12. - S. 474-480.

[7] Clark J., Demircan M., Kettas K. Europe: the EU AI Act's relationship with data protection law – key takeaways. - DLA Piper, Privacy Matters Blog, April 25, 2024

[8] Technological challenges of the introduction of artificial intelligence and machine learning in safety: the balance between innovation and responsibility/N. Sh. Kozlova, V. A. Dovgal, A. K. Dorgushaova, R. S. Kozlov//Bulletin of Adygea State University. Series: Natural-mathematical and technical sciences. – 2025. – № 1(356). - S. 55-69.

[9] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (Text with EEA relevance) // https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng (дата обращения: 17.03.2026)

[10] Titkov, D.I. Reznichenko, S.A. Revolution in information security audit: artificial intelligence/D.I. Titkov//Bulletin of Science, 2024. - VOL. 1. № 4(73). – С. 431-438.