УДК / UDC 340
DOI: 10.34076/22196838_2026_1_87

ПРАВОВОЕ РЕГУЛИРОВАНИЕ ФОРМИРОВАНИЯ И ИСПОЛЬЗОВАНИЯ РЕГИОНАЛЬНЫХ СОСТАВОВ ДАННЫХ В ПРОЦЕССЕ ОБУЧЕНИЯ МОДЕЛЕЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Олифиренко Артем Алексеевич
Магистрант Саратовской государственной юридической академии (Саратов), магистрант Саратовского государственного технического университета (Саратов), специалист по защите данных, ответственный за безопасность ИИ-систем ООО «Экосистема недвижимости „Метр квадратный“» (Москва), ORCID: 0000-0002-2186-281X, e-mail: panolifer@lamirtech.su.

В статье представлена обстоятельная научно-правовая интерпретация института региональных составов данных, формализованного в рамках экспериментального правового режима, предусмотренного федеральными законами № 123-ФЗ и № 233-ФЗ, регулирующими использование массивов обезличенной информации при обучении моделей искусственного интеллекта. Раскрывается сущностное содержание правового механизма формирования, обработки и последующего использования региональных составов данных в пределах государственной информационной инфраструктуры, действующей по модели технологического суверенитета. Анализируются дифференцированные требования к субъектам допуска к данным, включая критерии институциональной аффилиации, национальной юрисдикции, а также соблюдение императивов информационной безопасности и комплаенс-контроля. Аргументируется трансформация правового режима обучения моделей искусственного интеллекта в плоскость публично-правового регулирования с приоритетом обеспечения правомерности и прозрачности алгоритмических процессов в стратегически значимых отраслях. Обосновывается необходимость установления правовой презумпции ответственности операторов за внедрение и эксплуатацию региональных составов данных, включая меры по мониторингу и аудиту функционирования указанных моделей. Предложены концептуальные направления гармонизации законодательства в части категорирования региональных составов данных для обучения ИИ-моделей, способствующих укреплению цифрового суверенитета Российской Федерации.

Ключевые слова: региональные составы данных, системы искусственного интеллекта, правовой режим, обработка персональных данных, комплаенс-контроль, технологический суверенитет, информационная безопасность

Для цитирования: Олифиренко А. А. Правовое регулирование формирования и использования региональных составов данных в процессе обучения моделей искусственного интеллекта // Электронное приложение к «Российскому юридическому журналу». 2026. № 1. С. 87–97. DOI: https://doi.org/10.34076/22196838_2026_1_87.

LEGAL REGULATION OF THE FORMATION AND UTILIZATION OF REGIONAL DATA SETS IN THE PROCESS OF ARTIFICIAL INTELLIGENCE MODEL TRAINING

Olifirenko Artem
Master’s student, Saratov State Law Academy (Saratov), Master’s student, Yuri Gagarin State Technical University of Saratov (Saratov), Data Protection Specialist and AI Systems Security Officer, Ecosystem of Real Estate «Square Meter» LLC (Moscow), ORCID: 0000-0002-2186-281X, e-mail: panolifer@lamirtech.su.

The article presents a comprehensive legal and scholarly interpretation of the institution of regional data sets, formalized within the framework of the experimental legal regime established by Federal Laws No. 123-FZ and No. 233-FZ, which regulate the use

of anonymized information arrays for the training of artificial intelligence models. The substantive legal mechanism governing the formation, processing and subsequent use of regional data sets within the state information infrastructure-operating under the model of technological sovereignty is examined in detail. The study analyzes the differentiated requirements imposed on data access subjects, including criteria of institutional affiliation, national jurisdiction and compliance with imperatives of information security and regulatory oversight. The transformation of the legal regime governing AI-model training into the domain of public law regulation is substantiated, with emphasis on ensuring the lawfulness and transparency of algorithmic processes in strategically significant sectors. The necessity of establishing a legal presumption of operator liability for the implementation and exploitation of regional data sets is argued, including the introduction of monitoring and audit mechanisms over the functioning of the trained models. The article proposes conceptual directions for harmonizing existing legislation concerning the categorization of regional data sets used in AI-model training, aimed at reinforcing the digital sovereignty of the Russian Federation.

Key words: regional data sets, artificial intelligence systems, legal regime, personal data processing, compliance control, technological sovereignty, information security

For citation: Olifirenko A. (2026) Legal regulation of the formation and utilization of regional data sets in the process of artificial intelligence model training. In *Elektronnoe prilozhenie k «Rossiiskomu yuridicheskomu zhurnalu»*, no. 1, pp. 87–97, DOI: http://doi.org/10.34076/22196838_2026_1_87.

Современный этап развития технологий искусственного интеллекта характеризуется высокой степенью неопределенности и трансгрессией цифровых процессов, «выходящих за пределы возможностей традиционных правовых механизмов»[1] и ставящих под сомнение применимость устоявшихся конструкций нормативного регулирования в данной сфере. Указом Президента РФ от 10 октября 2019 г. № 490 «О развитии искусственного интеллекта в Российской Федерации» подчеркивается необходимость обеспечения «ускоренного внедрения ИИ в сферу государственного управления и экономики», что влечет за собой институциональные и регуляторные последствия. В связи с этим становится закономерным возникновение специальных правовых институтов, адаптированных под специфику технологической среды. Одним из таких институтов является экспериментальный правовой режим (далее – ЭПР), впервые введенный в российское законодательство Федеральным законом от 24 апреля 2020 г. № 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“» (далее – ФЗ № 123).

В соответствии со ст. 1 ФЗ № 123 экспериментальный правовой режим представляет собой особый нормативный порядок, установленный на ограниченный срок – пять лет с даты его введения (1 июля 2020 г.) – и реализуемый исключительно на территории города федерального значения Москвы. Его правовая конструкция направлена на создание условий для тестирования и внедрения инновационных решений в сфере искусственного интеллекта в специально определенных нормативных рамках. Именно с момента вступления данного Закона в силу началось формирование нормативной архитектуры, ориентированной на правовое сопровождение технологий ИИ в условиях экспериментального регулирования. Как подчеркивается в докладе ВШЭ, «именно на уровне города Москвы предусматривается… уникальный массив регулирования, обозначенный специальным федеральным законом, аналогов которому не существует в подавляющем большинстве иных субъектов РФ»[2]. Важным элементом данной нормы является введение специального реестра участников режима (ст. 5 ФЗ

[1] Правовые аспекты использования искусственного интеллекта: актуальные проблемы и возможные решения: доклад НИУ ВШЭ / В. Б. Наумов [и др.]. М.: Изд. дом ВШЭ, 2021. С. 2.

[2] Регулирование данных в Российской Федерации: текущее состояние, проблемы, перспективы: доклад НИУ ВШЭ / М. В. Якушев [и др.]. М.: Изд. дом ВШЭ, 2021. С. 14.

№ 123), включение в который осуществляется на основе строгих критериев, устанавливаемых уполномоченным органом исполнительной власти Москвы. Данный подход обеспечивает эффективный контроль за деятельностью субъектов, применяющих экспериментальное регулирование. Подобная система, как подчеркивается, позволяет поддерживать баланс интересов между «государством, высокотехнологичными компаниями-разработчиками, обществом, институтами развития и потребителями услуг»[1].

Федеральный законодатель определяет содержание экспериментального правового режима через категорию специального регулирования, отличного от общепринятого нормативного порядка (ст. 2 ФЗ № 123). Согласно законодательной дефиниции объектом правового регулирования выступают технологии искусственного интеллекта, понимаемые как «комплекс технологических решений», включающий программное обеспечение, ИКТ-инфраструктуру, а также алгоритмы обработки данных и принятия решений. Закон направлен на правовое обеспечение функционирования систем искусственного интеллекта в целом. Вместе с тем в рамках настоящего исследования акцент сделан исключительно на обучении ИИ-моделей как функциональной основе соответствующих систем, определяющей их способность к адаптивному и автономному принятию решений. Как справедливо подчеркивается в литературе, «машинное обучение лежит в основе системы принятия решений алгоритмами ИИ» и оказывает определяющее влияние на точность, воспроизводимость и правомерность соответствующих выводов[2].

Законодатель в ст. 3 ФЗ № 123 конкретизирует цели и задачи ЭПР, выделяя в качестве приоритетных направлений повышение эффективности государственного управления, стимулирование инновационной активности в экономической сфере и формирование комплексной правовой базы для регулирования отношений в области искусственного интеллекта. Центральными принципами ЭПР, закрепленными в указанной статье, становятся обязательная защита прав и свобод граждан, обеспечение безопасности личности и государства, а также недопущение дискриминации при использовании технологий ИИ. Данные принципы приобретают особую значимость в условиях обработки и анализа больших объемов персональных и обезличенных данных, используемых для обучения ИИ-моделей.

Следует особо отметить, что специфика автоматизированного анализа и использования персональных данных, в том числе в целях обучения ИИ, обусловила необходимость существенной корректировки федерального законодательства о персональных данных. В результате реализации поручений Президента РФ и федерального проекта «Нормативное регулирование цифровой среды» был принят Федеральный закон от 8 августа 2024 г. № 233-ФЗ[3] (далее – ФЗ № 233). Этим нормативным актом внесены изменения в ст. 6 и 10 Федерального закона от 27 июля 2006 г. № 152-ФЗ «О персональных данных» (далее – ФЗ № 152) и дополнения в ФЗ № 123. Нововведения направлены на создание современных правовых механизмов для безопасной и эффективной обработки данных, используемых при обучении алгоритмов искусственного интеллекта.

Центральной новеллой обновленного законодательства является введение на федеральном уровне института «региональных составов данных». Согласно положениям п. 3 ст. 13.1 ФЗ № 152 региональные составы данных представляют собой специальные массивы обезличенной информации, формируемые путем обработки персональных сведений, полученных от государственных и муниципальных органов власти и иных подведомственных организаций. При этом обезличивание данных осуществляется таким образом, чтобы исключить прямую или косвенную идентификацию субъекта

[1] Правовое регулирование технологий искусственного интеллекта: опыт России и Франции / М. А. Егорова [и др.] // Евразийская адвокатура. 2020. № 4. С. 80.

[2] *Наумов В. Б., Тытюк Е. В.* Правовые проблемы машинного обучения // Образование и право. 2020. № 6. URL: https://cyberleninka.ru/article/n/pravovye-problemy-mashinnogo-obucheniya (дата обращения: 21.03.2025).

[3] Федеральный закон от 8 августа 2024 г. № 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"“».

данных без привлечения дополнительной информации, доступ к которой жестко регламентируется законодательно.

Статья 6.1 ФЗ № 123 определяет, что процедура создания региональных составов данных осуществляется под непосредственным контролем высших исполнительных органов государственной власти субъекта Российской Федерации. В Москве такую функцию выполняет Правительство города, наделенное полномочиями по утверждению условий и требований к формированию составов данных, а также по контролю за соблюдением всех этапов их создания, обработки и последующего использования (ч. 1 ст. 6.1 ФЗ № 123). Правительство Москвы также устанавливает перечень данных, подлежащих передаче, и сроки осуществления такой передачи от соответствующих операторов (ч. 2 ст. 6.1 ФЗ № 123).

Обезличивание персональных данных представляет собой сложный технологический процесс, четко регламентированный в приказе Роскомнадзора № 966[1]. Среди используемых методов, как уже указывалось, особое место занимают введение идентификаторов (псевдонимизация), декомпозиция, семантическое преобразование и перемешивание данных.

Псевдонимизация в данном контексте представляет собой замену персонально идентифицирующих атрибутов (например, фамилии, имени, отчества, адреса, номера телефона или иного уникального идентификатора) на специально сгенерированные и неинформативные коды (псевдонимы), которые не позволяют идентифицировать конкретное лицо без доступа к дополнительному ключу дешифровки, хранящемуся в изолированной защищенной среде. В отличие от полной деперсонализации, псевдонимизация сохраняет возможность обратной идентификации в строго контролируемых условиях, что делает ее применимой в тех случаях, когда требуется сохранить аналитическую ценность массива данных без нарушения прав субъектов персональных данных.

Декомпозиция данных представляет собой технологию, согласно которой исходный массив персональной информации разбивается на отдельные фрагменты таким образом, что становится практически невозможным восстановить исходные персональные сведения субъекта.

Особые требования закон устанавливает и к порядку допуска операторов и организаций к региональным составам данных для обучения ИИ-моделей. В соответствии с п. 7 ст. 13.1 ФЗ № 152 допускаются исключительно российские юридические лица и граждане РФ, соответствующие строго определенным законодательным требованиям. Данные ограничения носят фундаментальный характер, поскольку обучение ИИ-моделей опирается на масштабные массивы данных, в том числе содержащие сведения о социально-экономических процессах в конкретных регионах Российской Федерации, что потенциально может повлиять на стратегические сферы государственного управления, безопасности и обороны.

Ключевым условием доступа операторов и иных субъектов к составам данных выступает подтверждение наличия сведений об операторе в реестре операторов персональных данных, который ведется в соответствии со ст. 22 ФЗ № 152. Данный реестр выполняет функцию публичного контроля за деятельностью субъектов обработки данных и служит одним из базовых инструментов обеспечения прозрачности обращения с персональной информацией. При этом «машинное обучение… требует создания человеком самой модели – вычислительной системы», а также снабжения ее «огромными объемами обучающих данных, анализ и интерпретация которых позволяет… улучшать ее»[2]. Современные архитектуры (глубокое обучение, включая сверточные нейронные сети, рекуррентные нейронные сети, трансформеры и их гибридные формы), применяемые в прикладных задачах, демонстрируют способность «улавливать закономерности в сложных генеративных паттернах»[3], что делает их функцио-

[1] Приказ Роскомнадзора от 5 сентября 2013 г. № 966 «Об утверждении требований и методов по обезличиванию персональных данных».

[2] *Рожкова М. А.* Машинное обучение как метод применения искусственного интеллекта: суть технологии и обзор основных правовых проблем // Machine learning as a method of applying artificial intelligence: legal issues // Закон.ру. 2024. 18 марта. URL: https://zakon.ru/blog/2024/03/18/mashinnoe_obuchenie_kak_metod_primeneniya_iskusstvennogo_intellekta_sut_tehnologii_i_obzor_osnovnyh_ (дата обращения: 21.03.2025).

[3] *Дюльдин Е. В., Зайцев К. С.* Применение глубокого обучения для выявления и классификации DGA доменов // International Journal of Open Information Technologies. 2022. № 8. URL: https://cyberleninka.ru/

нальным ядром интеллектуальных решений. В связи с этим запись субъекта в реестре операторов приобретает значение формального свидетельства соблюдения законодательства о персональных данных, а также подтверждения квалификации и ответственности за корректное использование информации в процессе обучения ИИ-моделей.

Вторым непременным критерием доступа является нахождение юридического лица под контролем Российской Федерации, субъекта РФ или муниципального образования либо российских граждан, не имеющих гражданства другого государства. Под контролем в данном случае понимается как наличие возможности определять решения юридического лица в силу права распоряжаться более чем 50 % голосующих акций или долей в уставном капитале данного субъекта хозяйственной деятельности. Эта норма направлена на исключение влияния иностранных структур на процесс обучения и использования ИИ-моделей, минимизацию рисков вывода чувствительных данных за пределы юрисдикции Российской Федерации и предотвращение возможного внешнего вмешательства в государственные цифровые системы, функционирующие на основе данных моделей.

Особое внимание законодателем уделяется вопросам безопасности и добросовестности лиц, получающих доступ к региональным составам данных. Отсутствие в Едином государственном реестре юридических лиц записей о недостоверности сведений о юридическом лице, а также непричастность руководящего состава организации к экстремистской или террористической деятельности подтверждаются в рамках предварительной проверки. Аналогичные требования предъявляются к гражданам РФ: отсутствие гражданства иностранного государства, отсутствие неснятой или непогашенной судимости за преступления, связанные с нарушением порядка обработки данных (включая ст. 183, 272, 273, 274.1, 283 и 283.1 УК РФ), а также отсутствие привлечения к уголовной ответственности по этим статьям за последние пять лет. Отметим, что перечень преступлений, закрепленный в подп. «ж» ч. 7 ст. 13.1 ФЗ № 152, не был приведен в соответствие с последующими изменениями уголовного законодательства: в УК РФ Федеральным законом от 30 ноября 2024 г.[1] была включена ст. 272.1, предусматривающая ответственность за незаконные использование и (или) передачу, сбор и (или) хранение компьютерной информации, содержащей персональные данные, однако в диспозиции подп. «ж» ч. 7 ст. 13.1 ФЗ № 152 данная статья прямо не упомянута. Следовательно, ст. 272.1 УК РФ содержательно относится к числу составов, непосредственно связанных с неправомерным обращением с персональными данными, однако формально в установленный законодателем перечень оснований для соответствующей проверки в рамках ст. 13.1 ФЗ № 152 в настоящее время не включена.

Вместе с тем доступ к таким данным регулируется дополнительными техническими и организационными мерами защиты. Как закреплено в п. 10 ст. 13.1 ФЗ № 152, обработка составов данных допускается исключительно в пределах государственной информационной системы уполномоченного органа в сфере регулирования информационных технологий. При этом все операции с данными, включая обучение ИИ-моделей, должны производиться непосредственно в инфраструктуре данной системы, без возможности извлечения или передачи составов данных за пределы защищенного периметра.

В постановлении Правительства Москвы от 3 декабря 2020 г. № 2134-ПП[2] отдельно прописан порядок тестирования, внедрения и использования технологий ИИ (после принятия ФЗ № 233), разработанных на основе региональных составов данных. Указанный акт требует от разработчиков обязательного проведения комплексной проверки эффективности и безопасности моделей, включая наличие механизмов аварийного отключения при обнаружении аномалий и других технических сбоев, способных привести к неправильному принятию решений и причинению ущерба. Для полноценного внедрения ИИ в государственное и муниципальное управление не-

article/n/primenenie-glubokogo-obucheniya-dlya-vyyavleniya-i-klassifikatsii-dga-domenov (дата обращения: 21.03.2025).

[1] Федеральный закон от 30 ноября 2024 г. № 421-ФЗ «О внесении изменений в Уголовный кодекс Российской Федерации».

[2] Постановление Правительства Москвы от 3 декабря 2020 г. № 2134-ПП «Об утверждении положения о проведении эксперимента по развитию технологий искусственного интеллекта в городе Москве».

обходимо «создание комплексной нормативной базы, которая бы регулировала вопросы ответственности при сбоях в работе систем ИИ, а также вопросы приватности и защиты персональных данных»[1].

В табл. 1 визуализирован предусмотренный ФЗ № 233 процесс сбора региональных данных и дальнейшего их использования для обучения ИИ-моделей.

Таблица 1

Схема процесса сбора, обработки и использования региональных составов данных по ФЗ № 233

Этап процесса	Описание процесса	Нормативно-правовое регулирование
1. Сбор персональных данных от региональных операторов	Персональные данные собираются государственными и муниципальными органами Москвы и передаются в региональную информационную систему согласно требованиям, установленным высшим исполнительным органом субъекта РФ (города Москвы). Перечень и сроки передачи данных заранее определяются уполномоченным органом	Часть 2 ст. 6.1 ФЗ № 123, ч. 2 ст. 13.1 ФЗ № 152
2. Предварительная проверка качества данных	Проверка качества, полноты и структурирование персональных данных перед их передачей. Включает отбор только релевантных данных согласно предварительно сформулированным требованиям	Часть 3 ст. 6.1 ФЗ № 123
3. Обезличивание персональных данных	Данные подвергаются процедуре обезличивания по методикам, установленным высшим исполнительным органом Москвы и Правительством РФ. Контроль за соблюдением осуществляют Роскомнадзор и ФСТЭК России. Методы включают введение идентификаторов, декомпозицию и перемешивание данных для исключения идентификации субъектов	Часть 5 ст. 6.1 ФЗ № 123, ч. 3 ст. 13.1 ФЗ № 152
4. Формирование региональных составов данных	После обезличивания персональные данные интегрируются в единые региональные составы данных. Данные группируются по определенным признакам (сферам, секторам, типам данных) для упрощения последующей аналитической обработки и использования для обучения ИИ-моделей	Часть 6 ст. 6.1 ФЗ № 123, ч. 5 ст. 13.1 ФЗ № 152
5. Предоставление доступа к региональным составам данных	Доступ к данным предоставляется только уполномоченным пользователям, прошедшим проверку по критериям закона: государственные органы, муниципальные организации, а также российские юридические лица и граждане РФ, имеющие подтверждение соответствия требованиям закона. Проверка пользователей проводится по специальной процедуре	Части 7, 8 ст. 6.1 ФЗ № 12, чч. 6, 7 ст. 13.1 ФЗ № 152
6. Обучение моделей искусственного интеллекта на региональных составах данных	Региональные составы данных используются для обучения ИИ-моделей строго внутри региональной информационной системы. Данные загружаются в среды разработки и тренировки моделей, которые размещены внутри защищенного периметра системы	Часть 10 ст. 6.1 ФЗ № 123

[1] *Брычеев А. С.* Государственное управление и искусственный интеллект: новые возможности для малого и среднего бизнеса // Вестник евразийской науки. 2024. Т. 16. № s5. URL: https://esj.today/PDF/31FAVN524.pdf (дата обращения: 21.03.2025).

Окончание табл. 1

7. Мониторинг и контроль обработки данных и результатов обучения моделей ИИ	Постоянный мониторинг процесса обработки данных и обучения ИИ-моделей на предмет соответствия законодательным требованиям и предотвращения потенциального вреда гражданам и государству. Вводится запрет на обработку данных, результаты которых могут причинить вред национальной безопасности и общественным интересам	Части 11, 12 ст. 6.1 ФЗ № 123, ч. 11 ст. 13.1 ФЗ № 152

В табл. 2 приведен пример нормативно-технологической реализации ИИ-моделей с использованием региональных составов данных в условиях экспериментального правового режима. Представленный подход отражает архитектурный сценарий, выстроенный в логике MLSecOps[1] как комплексной модели правового, организационного и инфраструктурного сопровождения жизненного цикла ИИ-модели. В рамках данного подхода обеспечивается не только технологическая последовательность обработки данных и обучения модели, но и правовая верификация допустимости соответствующих операций, контроль соблюдения требований законодательства о персональных данных, а также минимизация рисков неправомерного использования данных, повторной идентификации субъектов персональных данных и возникновения иных юридически значимых последствий. Это позволяет последовательно отразить ключевые стадии такого процесса, включая загрузку регионального состава данных в защищенную среду государственной информационной системы, проверку правомерности и допустимости его использования, обучение и валидацию ИИ-модели, последующее тестирование, а также внедрение алгоритмических решений в пределах государственного информационного периметра.

Таблица 2

Этапы использования региональных составов данных для обучения ИИ-моделей в пределах государственного информационного периметра

Этап	Содержание этапа
1. Формирование и загрузка регионального состава данных	Региональный состав данных, сформированный по инициативе уполномоченного органа в соответствии с чч. 1–3 ст. 13.1 ФЗ № 152, подлежит загрузке в защищенную среду государственной информационной системы субъекта РФ. Передача такого состава допускается только после прохождения процедуры обезличивания, отвечающей требованиям, утвержденным Правительством РФ и согласованным с уполномоченными органами в сфере безопасности
2. Юридически допустимая обработка состава данных в инфраструктуре государственной информационной системы (ГИС)	В пределах инфраструктуры ГИС осуществляется юридически допустимая обработка состава данных, включающая верификацию источника происхождения данных, степени агрегированности информации и ее пригодности для обучения ИИ-моделей. На данной стадии формируется вывод о правомерности целей обработки и об отсутствии недопустимых рисков повторной идентификации субъектов персональных данных
3. Правовая валидация и регистрация обучения ИИ-модели	ИИ-модель в процессе обучения фиксируется в журнале операций в рамках государственной информационной системы и подлежит предварительной правовой валидации. Такая валидация осуществляется на предмет соответствия режима обработки установленным ограничениям, целям использования модели и допустимости применения результатов обучения в практической деятельности. Все действия подлежат обязательному протоколированию

[1] Space ISAC AI/ML Community. Machine Learning Security Operations – MLSecOps // Space ISAC White Paper. 2023. 15 p. URL: https://spaceisac.org/wp-content/uploads/2023/08/Space-ISAC-MLSecOps-White-Paper-08.04.2023.pdf (дата обращения: 21.03.2025).

Окончание табл. 2

4. Допуск субъекта к использованию состава данных для обучения	Использование состава данных допускается только субъектами, прошедшими проверку на соответствие критериям, установленным ч. 7 ст. 13.1 ФЗ № 152. Допуск оформляется в виде зарегистрированного решения о начале процедуры обучения с указанием цели обработки, предполагаемого алгоритма и модели, подлежащей использованию
5. Тестирование и контрольная валидация модели	По завершении процедуры обучения модель может быть протестирована в пределах контура государственной информационной системы исключительно на обезличенных контрольных выборках. Результаты тестирования подлежат анализу на предмет потенциальной дискриминации (под которой в данном случае понимается необоснованное различие в результатах обработки данных, способное повлечь ущемление прав и законных интересов отдельных категорий субъектов), нарушения прав субъектов данных либо возникновения угроз национальной безопасности
6. Внедрение модели и последующий мониторинг	При подтверждении соответствия модели установленным требованиям публично-правового регулирования она может быть внедрена в продуктивный контур информационной системы исключительно для выполнения задач, определенных стратегическими документами или иными нормативными актами. Дальнейшее функционирование модели должно сопровождаться постоянным аудитом, мониторингом и юридической подотчетностью оператора

Можно заключить, что внесение изменений в нормативную конструкцию ФЗ № 123 и ФЗ № 152 посредством ФЗ № 233 инициировало трансформацию правового режима обучения ИИ-моделей, легитимировав на уровне законодательства новый правовой институт – региональные составы данных. Указанный институт не только создал формализованную основу для централизации и унификации работы с обезличенной информацией, но и модифицировал основания правомерности использования данных в процессе обучения ИИ-моделей на территории Российской Федерации.

Ключевой особенностью института региональных составов данных выступает нормативное закрепление специального режима обработки обезличенной информации, предназначенной для использования в целях машинного обучения. В отличие от ранее действовавших механизмов, основанных преимущественно на диспозитивной модели обращения с данными в рамках ФЗ № 152, новый режим вводит вертикаль полномочий в части как формирования, так и последующего использования обезличенных составов данных субъектами, допущенными к их эксплуатации. Тем самым доступ операторов (ИТ-компаний) к данным ограничивается рамками публично-правового регулирования, включающего многоуровневую проверку правового статуса субъекта, а также его соответствие критериям доверия, установленным в п. 7 ст. 13.1 ФЗ № 152.

Изменению подверглась и природа нормативного контроля за стадией обучения: если ранее предмет регулирования охватывал лишь факт передачи массивов персональных или обезличенных данных на основе гражданско-правовых соглашений, то после вступления в силу ФЗ № 233 под нормативное управление подпадает весь жизненный цикл данных: от их первоначального сбора и процедур обезличивания до загрузки в ИИ-модель.

На этом фоне закономерным представляется смещение вектора правового регулирования обучения ИИ-моделей из сферы преимущественно частноправовых конструкций – ориентированных на режим согласия, контрактного доступа к данным и локальной ответственности разработчика – в плоскость публично-правового регулирования, сопряженного с превалированием критериев правомерности, транспарентности и контролируемости алгоритмических процессов. Такая трансформация правовой природы обучения ИИ обусловлена не только необходимостью защиты персональных данных, но и стратегическим значением алгоритмически автоматизированных решений в таких секторах, как государственное управление, оборонная

безопасность, здравоохранение и социальная политика. Установление институциональной нормативной инфраструктуры, обеспечивающей постоянный надзор за параметрами, логикой и результатами алгоритмов, становится не факультативным элементом, а правовым императивом, закрепляющим подотчетность цифровых решений интересам общества и государства.

Институт региональных составов данных, таким образом, смещает акцент с договорной модели обработки данных в сторону императивного публично-правового регулирования. Это выражается в закреплении обязательных требований к безопасности, институциональной аффилированности оператора и допустимости осуществления обработки данных исключительно в рамках публичных процедур. Факт правомерности обучения ИИ-моделей напрямую зависит от соответствия субъекта установленным нормативным критериям и его включенности в систему государственного надзора.

Кроме того, данный институт трансформирует структуру правовой легитимации результатов обучения. Согласно п. 11 ст. 13.1 ФЗ № 152 прямо запрещается использование результатов обработки, если их применение способно причинить ущерб национальным интересам, обороноспособности, безопасности либо иным охраняемым законом ценностям. Так, обученные ИИ-модели приобретают правовой статус объектов, потенциально затрагивающих сферу публичного порядка, что, в свою очередь, порождает обязанность по обеспечению институционального мониторинга последствий их функционирования.

С учетом публично-правовой природы региональных составов данных, закрепленной в ст. 13.1 ФЗ № 152, использование таких данных в целях обучения ИИ-моделей должно сопровождаться надлежащей ответственностью за нарушение установленного режима их обработки. В то время как законодатель обеспечил многоступенчатую систему допуска пользователей к государственной информационной системе, содержащей составы данных (ч. 7 ст. 13.1), а также закрепил запрет на использование таких массивов вне инфраструктуры, находящейся под контролем уполномоченного органа (ч. 10 ст. 13.1), административное законодательство не предусматривает специальной нормы, определяющей принудительное исполнение этих положений посредством санкционного механизма.

Представляется необходимым включить в недавно измененную ст. 13.11 КоАП РФ новую часть, устанавливающую ответственность за нарушение режима использования региональных составов данных, сформированных на основе обезличивания персональных данных. В частности, предлагается закрепить административную ответственность за обработку таких данных вне государственной информационной системы, нарушение установленного процедурного порядка допуска, предоставление доступа третьим лицам, не соответствующим требованиям п. 7 ст. 13.1 ФЗ № 152, а также за отсутствие документируемого мониторинга операций по доступу и обработке в рамках обучающих контуров ИИ-моделей. Диспозиция предлагаемой части должна быть сформулирована по аналогии с чч. 12–17 ст. 13.11 КоАП РФ, предусматривающими ответственность за массовые нарушения при обработке персональных данных, однако с учетом специфики и публично-правового статуса региональных составов данных. Так, санкции за нарушения должны быть установлены на уровне от 300 тыс. до 600 тыс. руб. для должностных лиц и от 3 млн до 10 млн руб. для юридических лиц. Это позволит ликвидировать имеющийся нормативный пробел и повысить гарантии правомерности и прозрачности использования стратегически значимых данных при обучении ИИ-моделей.

Необходимо нормативное закрепление категорирования региональных составов данных, направленное на дифференцированный подход к допуску субъектов к таким данным в пределах государственной информационной системы, находящейся под контролем уполномоченного органа. Действующая редакция ст. 13.1 ФЗ № 152 исключает публичный доступ к таким составам данных, вследствие чего категорирование не должно упрощать внешний доступ, а призвано обеспечить внутреннюю нормативную дифференциацию в рамках ГИС.

Предлагается закрепить соответствующее категорирование на уровне ведомственного приказа Минцифры России, поскольку именно на данный орган возложена обязанность нормативного регулирования порядка формирования составов данных (ч. 5

ст. 13.1) и установления регламентов допуска к ним (ч. 6 ст. 13.1). Такое категорирование должно учитывать специфические особенности различных составов данных, обусловленные степенью их чувствительности, уровнем детализации и агрегированности информации, а также функциональным назначением и потенциальными рисками повторной идентификации субъектов.

Предлагаемая структура категорирования должна базироваться на четких и юридически значимых критериях (см. табл. 3).

Таблица 3

Предлагаемая модель категорирования региональных составов данных

Категория доступа	Критерии категорирования данных	Субъекты, имеющие право доступа
Категория I (высокочувствительный доступ)	высокая источниковая чувствительность (сферы здравоохранения, безопасности, социальной поддержки); низкая степень обобщенности и высокая детализация данных; высокий риск повторной идентификации через другие источники; функциональное назначение ИИ-моделей для критически важных государственных решений; высокий уровень допуска субъектов, строгая проверка и подконтрольность РФ	Государственные органы и их подведомственные учреждения
Категория II (контролируемый научно-исследовательский доступ)	средняя источниковая чувствительность данных; средняя степень обобщенности и детализации данных; умеренный риск повторной идентификации; использование ИИ-моделей для научно-исследовательских и аналитических целей; строгий контроль субъектов с обязательной регистрацией проектов и мониторингом моделей	Научные и аналитические организации, соответствующие требованиям ч. 7 ст. 13.1 ФЗ № 152, прошедшие комплексную проверку, при условии постоянного мониторинга и валидации ИИ-моделей
Категория III (условно обособленная аналитика)	низкая источниковая чувствительность данных (агрегированные массивы без чувствительной информации); высокая степень обобщенности и минимальная детализация данных; низкий риск повторной идентификации; аналитическое и статистическое функциональное назначение; упрощенная процедура допуска, но полная изоляция данных	Субъекты, соответствующие требованиям ч. 7 ст. 13.1 ФЗ № 152, доступ возможен только в рамках защищенной инфраструктуры ГИС, с запретом внешнего распространения результатов

Интеграция института региональных составов данных в нормативную архитектуру обращения с данными, используемыми в процессе обучения ИИ-моделей, обеспечила формирование внутренне согласованного, правомерного и предсказуемого режима, в рамках которого достигается институциональный баланс интересов государства, общества и бизнеса. С учетом специфики применения ИИ-моделей в сфере государственного управления, социальной политики и обеспечения национальной безопасности институционализация соответствующего механизма является не только легитимной, но и объективно необходимой в условиях усложняющейся цифровой реальности. В этом контексте ФЗ № 123 должен рассматриваться как временный, но системообразующий элемент, правовые конструкции и результаты применения которого подлежат последующей инкорпорации в систему общего регулирования обращения с данными и обучения ИИ-моделей при одновременном сохранении на федеральном уровне института региональных составов данных.

Список литературы

Брычеев А. С. Государственное управление и искусственный интеллект: новые возможности для малого и среднего бизнеса // Вестник евразийской науки. 2024. Т. 16. № s5. URL: https://esj.today/PDF/31FAVN524.pdf (дата обращения: 21.03.2025).

Дюльдин Е. В., Зайцев К. С. Применение глубокого обучения для выявления и классификации DGA доменов // International Journal of Open Information Technologies. 2022. № 8. С. 3–10. URL: https://cyberleninka.ru/article/n/primenenie-glubokogo-obucheniya-dlya-vyyavleniya-i-klassifikatsii-dga-domenov (дата обращения: 24.03.2025).

Наумов В. Б., Тытюк Е. В. Правовые проблемы машинного обучения // Образование и право. 2020. № 6. С. 219–231. URL: https://cyberleninka.ru/article/n/pravovye-problemy-mashinnogo-obucheniya (дата обращения: 21.03.2025).

Правовое регулирование технологий искусственного интеллекта: опыт России и Франции / М. А. Егорова [и др.] // Евразийская адвокатура. 2020. № 4. С. 79–82.

Правовые аспекты использования искусственного интеллекта: актуальные проблемы и возможные решения: доклад НИУ ВШЭ / В. Б. Наумов [и др.]. М.: Изд. дом ВШЭ, 2021. 42 с.

Регулирование данных в Российской Федерации: текущее состояние, проблемы, перспективы: доклад НИУ ВШЭ / М. В. Якушев [и др.]. М.: Изд. дом ВШЭ, 2021. 33 с.

Рожкова М. А. Машинное обучение как метод применения искусственного интеллекта: суть технологии и обзор основных правовых проблем // Machine learning as a method of applying artificial intelligence: legal issues // Закон.ру. 2024. 18 марта. URL: https://zakon.ru/blog/2024/03/18/mashinnoe_obuchenie_kak_metod_primeneniya_iskusstvennogo_intellekta_sut_tehnologii_i_obzor_osnovnyh_ (дата обращения: 21.03.2025).

References

Brycheev A. S. (2024) Gosudarstvennoe upravlenie i iskusstvennyi intellekt: novye vozmozhnosti dlya malogo i srednego biznesa [State Governance and Artificial Intelligence: New Opportunities for Small and Medium Enterprises]. In *Vestnik evraziiskoi nauki*, vol. 16, no. s5, available at: https://esj.today/PDF/31FAVN524.pdf (accessed: 21.03.2025).

Dyuldin E. V., Zaitsev K. S. (2022) Primenenie glubokogo obucheniya dlya vyyavleniya i klassifikatsii DGA domenov [Application of Deep Learning for Detection and Classification of DGA Domains]. In *International Journal of Open Information Technologies*, no. 8, pp. 3–10, available at: https://cyberleninka.ru/article/n/primenenie-glubokogo-obucheniya-dlya-vyyavleniya-i-klassifikatsii-dga-domenov (accessed: 24.03.2025).

Egorova M. A., et al. (2020) Pravovoe regulirovanie tekhnologii iskusstvennogo intellekta: opyt Rossii i Frantsii [Legal Regulation of Artificial Intelligence Technologies: Experience of Russia and France]. In *Evraziiskaya advokatura*, no. 4, pp. 79–82.

Naumov V. B., et al. (2021) *Pravovye aspekty ispol'zovaniya iskusstvennogo intellekta: aktual'nye problemy i vozmozhnye resheniya: doklad NIU VSHE* [Legal Aspects of Artificial Intelligence Use: Current Problems and Possible Solutions: report of the National Research University Higher School of Economics]. Moscow, Izdatel'skii dom VSHE, 42 p.

Naumov V. B., Tytyuk E. V. (2020) Pravovye problemy mashinnogo obucheniya [Legal issues of machine learning]. In *Obrazovanie i pravo*, no. 6, pp. 219–231, available at: https://cyberleninka.ru/article/n/pravovye-problemy-mashinnogo-obucheniya (accessed: 21.03.2025).

Rozhkova M. A. (2024) Mashinnoe obuchenie kak metod primeneniya iskusstvennogo intellekta: sut' tekhnologii i obzor osnovnykh pravovykh problem [Machine Learning as a Method of Applying Artificial Intelligence: Legal Issues]. In *Zakon.ru*, 18 March, available at: https://zakon.ru/blog/2024/03/18/mashinnoe_obuchenie_kak_metod_primeneniya_iskusstvennogo_intellekta_sut_tehnologii_i_obzor_osnovnyh_ (accessed: 21.03.2025).

Yakushev M. V., et al. (2021) *Regulirovanie dannykh v Rossiiskoi Federatsii: tekushchee sostoyanie, problemy, perspektivy: doklad NIU VSHE* [Regulation of Data in the Russian Federation: Current State, Problems, Prospects: report of the National Research University Higher School of Economics]. Moscow, Izdatel'skii dom VSHE, 33 p.

Дата поступления рукописи в редакцию: 30.11.2025
Дата принятия рукописи в печать: 27.01.2026