

УДК 004.04

АНАЛИЗ НОРМАТИВНОГО РЕГУЛИРОВАНИЯ В ОБЛАСТИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА И ПРЕДЛОЖЕНИЯ ПО ЕГО РАЗВИТИЮ

Зацаринный А. А., *Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Москва, Россия, Вавилова 44/2, e-mail: frccsc@frccsc.ru*

Гаврилов В. Е., *Федеральный исследовательский центр «Информатика и управление» Российской академии наук, Москва, Россия, Вавилова 44/2*

В статье рассмотрены предпосылки и факторы, определяющие актуальность нормативного регулирования в области ИИ. Представлен анализ понятийного аппарата и основных положений принятого Федерального закона о поддержке развития технологий искусственного интеллекта в РФ с учетом накопленного опыта в области разработки и использования систем искусственного интеллекта. Рассмотрены проблемы технической реализуемости положений закона и достаточности соответствующего нормативно-методического обеспечения. Сформулированы предложения по развитию технического регулирования при реализации положений закона.

Ключевые слова: нормативно-правовое регулирование, большая фундаментальная модель искусственного интеллекта, информационная безопасность, комплексная безопасность, функциональная безопасность, системы с искусственным интеллектом, техническая защита информации.

ANALYSIS OF THE DRAFT FEDERAL LAW "ON SUPPORT FOR THE DEVELOPMENT OF ARTIFICIAL INTELLIGENCE TECHNOLOGIES IN THE RUSSIAN FEDERATION" AND ITS IMPACT ON THE DEVELOPMENT OF REGULATORY FRAMEWORK IN THE FIELD OF ARTIFICIAL INTELLIGENCE

Zatsarinny A. A., *Federal Research Center "Informatics and Control" of the Russian Academy of Sciences, Moscow, Russia, Vavilova 44/2, e-mail: frccsc@frccsc.ru*

Gavrilov V. E., *Federal Research Center "Informatics and Control" of the Russian Academy of Sciences, Moscow, Russia, Vavilova 44/2*

The article analyzes the conceptual apparatus of the draft Federal Law and some of its provisions, taking into account the accumulated experience in the development and use of artificial intelligence systems, considers the problems of technical feasibility of the provisions of the draft law and the sufficiency of the relevant regulatory and methodological support. Some proposals for the development of technical regulation in the implementation of the provisions of the draft law are formulated.

Keywords: legal regulation, large fundamental model of artificial intelligence, information security, integrated security, functional safety, systems with artificial intelligence, technical protection of information.

Введение

Развитие технологий ИИ (ТИИ) определено в рамках первого приоритета новой Стратегии научно-технологического развития России [1]. Эти технологии сегодня определяют будущее многих отраслей и занимают центральное место среди ключевых направлений мирового технологического прогресса. Президент России В.В. Путин на совещании в Крыму поручил Правительству вместе с регионами сформировать национальный план внедрения ИИ и добиться повсеместного использования этих технологий к 2030 году во всех областях, включая экономику и образование [2]. Важно, что он отметил необходимость развития отечественных фундаментальных моделей искусственного интеллекта, которые должны быть конкурентоспособными в мировом масштабе. Вместе с тем, понятно, что успешное выполнение поставленных Президентом задач невозможно без нормативного регулирования применения ТИИ, особенно применительно к сфере информационной безопасности систем с искусственным интеллектом (СИИ). В этой области уже накопился ряд системных проблем при создании СИИ в различных сферах, включая оборону, безопасность и правопорядок. Многие из них рассмотрены в [3,4,5].

Отметим, что к настоящему времени в Российской Федерации принято около 160 национальных и

предварительных стандартов в области разработки и использования систем искусственного интеллекта как общего характера, так и специализированных отраслевых. Действуют отраслевые методические документы в этой сфере, однако все эти документы направлены на регулирование сугубо технических и (или) технологических вопросов разработки и использования СИИ и их компонентов.

В этом году Правительством РФ сделан важный шаг по нормативному регулированию отношений в процессе разработки и применения ТИИ на федеральном уровне. Так, 18 марта 2026 года опубликован проект Федерального Закона «Об основах государственного регулирования применения технологий искусственного интеллекта» [6], разработанный в развитие положений утвержденной указом Президента РФ Национальной стратегии развития искусственного интеллекта, которая определяет основные цели, принципы и направления создания комплексной системы нормативно-правового регулирования общественных отношений, связанных с развитием и использованием технологий искусственного интеллекта в России [7]. В проекте Федерального закона впервые предпринята попытка законодательно отрегулировать отношения сторон-участников процессов разработки, создания и использования ТИИ как основного компонента многих СИИ.

Публикация проекта Федерального закона «Об искусственном интеллекте в Российской Федерации и внесении изменений в отдельные законодательные акты Российской Федерации» вызвала живой интерес в ИТ-сообществе и по итогам общественного обсуждения документ претерпел радикальные изменения [8]. Далее в июле 2026 г. он в ускоренном режиме был принят в Государственной Думе (1-е чтение - 7 июля, 2-е и 3-е — 8 июля) и 17 июля 2026 г. одобрен на 617-м заседании Совета Федерации [9] и подписан Президентом России [10]. Определено, что документ вступит в силу поэтапно: с 1 сентября 2026 г. - основные положения, с 1 марта 2027 г. - нормы о статусах моделей, маркировке и интеллектуальной собственности. Интересно, что для уже работающих систем, использующих ТИИ, предусмотрен пятилетний переходный период до 1 сентября 2032 года.

Таким образом, закон принят и поэтому представляет интерес анализ его влияния на дальнейшее развитие ИИ в России.

1. Предпосылки и факторы, определяющие актуальность нормативного регулирования в области ИИ.

Как уже отмечено выше, предпосылки нормативного регулирования применения ИИ на федеральном уровне созрели уже несколько лет назад.

Сформулируем факторы, определяющие актуальность и необходимость нормативного регулирования применения ИИ. Они обусловлены прежде всего тем, что ТИИ обладают принципиально двойственной природой: с одной стороны, они обладают огромными, можно сказать – неисчерпаемыми, возможностями для повышения эффективности деятельности самых разных структур и процессов, с другой стороны, создают широкий спектр разнообразных угроз и рисков в обеспечении функционирования этих структур в соответствии с заданными правилами и регламентами. Выделим основные группы факторов, имеющих различный характер (технических, технологических, нравственно-этических и даже геополитических) и определяющих нормативное регулирование применения ТИИ как приоритетную проблему стратегического значения.

Первая группа состоит в необходимости обеспечения национальной безопасности государства, т.к. неконтролируемое применение и развитие ТИИ может быть использовано в целях дестабилизации политической и экономической обстановки за счет ведения кибершпионажа и гибридных войн. Достаточно упомянуть целый ряд нормативных актов администрации США, ориентированных на интенсивное технологическое развитие для обеспечения национальной безопасности и обороны страны, важное место в которых занимает ИИ (например, [11,12]). России нужен свой ИИ, полностью контролируемый обществом и государством, и способный противостоять потенциальным угрозам.

Вторая группа обусловлена трансграничным характером потенциальных угроз применения ТИИ, которые, по сути, не имеют государственных границ. Другими словами, существует объективная необходимость нормативного регулирования на международном уровне на основе разработки универсальных кодексов поведения разработчиков и гармонизации национальных стандартов безопасности применения ТИИ.

Третья группа факторов – огромные риски для обеспечения защиты персональных данных, т.к. с использованием ТИИ обрабатываются огромные объемы конфиденциальных данных, которые является привлекательной целью для совершения киберпреступлений. Нормативный контроль должен снизить риски утечки персональных данных, нарушения неприкосновенности частной жизни, несанкционированной слежки за физическими лицами, особенно при использовании систем биометрического распознавания.

Четвертая группа – проблема отсутствия юридической ответственности за ущерб, причиненный в результате применения систем с использованием ТИИ. Конкретно речь идет о необходимости нормативного механизма

разграничения и конкретизации ответственности между субъектами в области создания и применения ТИИ (заказчик, разработчик, оператор и пользователь).

Пятая группа обусловливается гуманитарными проблемами применения ТИИ, т.к. они обладают низким уровнем интерпретируемости (объяснимости) решений, принимаемых сложными нейросетями, которые зачастую создаются с использованием недоверенных аппаратно-программных средств (проблема «черного ящика»). Эти проблемы порождают множество нравственно-этических рисков, особенно при использовании ТИИ в системах критической информационной инфраструктуры, и требуют нормативного закрепления принципов прозрачности, объяснимости и обязательного аудита ИИ-алгоритмов перед их внедрением.

Таким образом, приведенные факторы указывают на то, что нормативное регулирование применения ИИ должно обеспечивать на основе риск-ориентированного подхода баланс между стимулированием развития ТИИ и обеспечением надежной защиты государства, общества и личности от новых киберугроз.

Рассмотрим, в какой мере принятый ФЗ [10] соответствует ожиданиям со стороны различных субъектов, участвующих в разработке и применении ТИИ, с учетом приведенных выше факторов.

2. Цель, направленность и основные положения ФЗ

В принятой редакции закон регулирует отношения, связанные с разработкой, внедрением, применением и использованием больших фундаментальных моделей (БФМ) искусственного интеллекта [10]. Под БФМ понимают программы, которые решают широкий круг задач и содержат не менее одного миллиарда параметров. Узкоспециализированные модели или прикладные сервисы на их основе под действие закона не подпадают.

Отметим, что в проекте ФЗ, опубликованном 18.03.2026 г. (статья 1. П.1) [6], целевая направленность закона представлялась более широко, а именно: «предметом нормативного регулирования являлись отношения, возникающие в связи с разработкой, внедрением, использованием искусственного интеллекта». Другими словами, предполагалось, что закон будет распространяться на все технологии ИИ.

В Законе определено шесть целей, суть которых кратко в следующем:

- создание условий (правовых, экономических и организационных) для развития БФМ в экономике, социальной сфере и госуправлении;
- поддержание исследований и инноваций в этой сфере;
- обеспечение безопасности личности, общества и государства, а также технологической независимости;
- повышение эффективности госуправления, развитие цифровой экономики и улучшение качества жизни;
- достижение технологического лидерства за счёт конкурентоспособных на мировом уровне моделей;
- развитие международного сотрудничества, в том числе для продвижения российских технологий за рубежом.

Закон вводит две категории БФМ – суверенная и национальная. Первая должна разрабатываться только российским юридическим лицом, при этом цикл ее создания, включая обучение, должен быть технически воспроизводим без критической зависимости от зарубежных компонентов. Национальную модель также должна разрабатывать российская компания, но при этом допускается использование иностранных компонентов и моделей с открытыми лицензиями. В обоих случаях хранение данных должно происходить в российских центрах обработки данных, принадлежащих российским юридическим лицам.

Закон носит рамочный характер. Это значит, что он задаёт общие рамки и базовые правила, а детальное регулирование (конкретные технические требования, перечни сфер, механизмы поддержки) будет осуществляться через подзаконные акты (постановления Правительства и другие нормативные документы).

3. Анализ ключевых положений и общая оценка

Главное в том, что сфера действия закона резко сужена – он теперь регулирует отношения, связанные с разработкой, внедрением и применением только больших фундаментальных моделей искусственного интеллекта. Такое, достаточно неожиданное изменение целевой направленности закона, привело к несоответствию его названия («О поддержке развития технологий искусственного интеллекта в РФ») содержанию, т.к. понятие технологии искусственного интеллекта существенно шире, чем большие фундаментальные модели [13]. В результате вне мер нормативного регулирования и соответственно государственной поддержки оказалось огромное число крайне важных для практического применения технологий искусственного интеллекта (экспертные системы, машинное зрение, промышленные IoT, распознавание аудио информации, ИИ-агенты и др.). Такая направленность закона уже вызвала неоднозначные оценки экспертов. В частности, Н.И. Касперская отметила, что выбор в качестве основного объекта регулирования именно больших фундаментальных моделей выглядит спорно, поскольку технологии ИИ

быстро меняются, например, сейчас на первое место выходят ИИ-агенты [14].

Кроме этого, отметим еще несколько принципиальных замечаний по существу закона.

1) В законе не определен перечень субъектов отношений, возникающих в связи с разработкой, внедрением, использованием и применением больших фундаментальных моделей искусственного интеллекта, их права и обязанности (ранее в статье 3 были перечислены: разработчик модели ИИ, оператор системы ИИ, владелец сервиса доступа и пользователь сервиса ИИ).

2) Из текста законопроекта [10] исключено понятие системы искусственного интеллекта. При этом на разработчика большой фундаментальной модели искусственного интеллекта как программы возлагается обязанность принимать организационные и технические меры по обеспечению ее безопасности (Ст.8, п.1). Заметим, что модель искусственного интеллекта (программа) функционирует в составе некоторой информационной системы – системы искусственного интеллекта, а меры обеспечения безопасности информации в большинстве случаев реализуются средствами программной аппаратной среды функционирования разработчиком системы. То же относится и к требованию о необходимости вести техническую документацию с описанием ключевых параметров и ограничений большой фундаментальной модели искусственного интеллекта в объеме, необходимом для оценки информационной безопасности ее применения (Ст.8, п.3). Такого рода документацию может представить только разработчик системы искусственного интеллекта с учетом решаемых задач и конкретных условий ее эксплуатации.

3) В законе недостаточно рассмотрены вопросы нормативно-методического обеспечения процессов разработки и применения ТИИ в части информационной безопасности и реализуемости положений законопроекта с учетом сложившейся ситуации в этой области [4,5].

4) Появилась отсылочная норма о возможности введения уточняющих норм при использовании больших фундаментальных моделей для целей обороны и безопасности государства, что представляется весьма важным, т.к. отражает общемировой тренд миграции апробированных в гражданской сфере технологий, включая ТИИ, в военную сферу.

5) В законе по-прежнему сохранена расширительная формулировка определения искусственного интеллекта Национальной стратегии [7], которая может охватывать широкий спектр технологий, включая базовые функции устройств (автокоррекцию текста, шумоподавление, обработку фото), что создаёт избыточные требования. Определение искусственного интеллекта включает в том числе, информационно-коммуникационную инфраструктуру, программное обеспечение, процессы и сервисы по обработке данных. При этом информационно-коммуникационная инфраструктура определяется в [15] как совокупность территориально распределённых государственных и корпоративных информационных систем, сетей связи, средств коммутации и управления информационными потоками т.е. выполняет разнородные задачи, не связанные, собственно, с искусственным интеллектом. Определение большой фундаментальной модели искусственного интеллекта сводит, по сути, все многообразие технологий искусственного интеллекта только к генеративной модели ИИС на основе машинного обучения.

Таким образом, принятый закон не в полной мере обеспечивает правовую поддержку целей, заявленных в Национальной стратегии [7].

4. Основные проблемы и предложения

4.1. Риск ориентированный подход

Одним из основных принципов регулирования отношений в сфере применения больших фундаментальных моделей искусственного интеллекта анонсирован риск-ориентированный подход - установление требований к применению больших фундаментальных моделей искусственного интеллекта с учетом вероятности рисков причинения вреда (ущерба) и масштабам причиняемого ущерба. Однако дальнейшего развития этот принцип в тексте законопроекта не получил. Представляется целесообразным определить классификацию по уровню потенциального ущерба при некорректном функционировании моделей: неприемлемый, высокий, средний и незначительный. В последующих подзаконных актах и нормативно-методических документах этот подход может быть закреплён с предъявлением дифференцированных требований безопасности СИИ. Такой подход применяется, в частности, в EU AI Act [16]. В связи с изложенным, предлагается:

- для государственных информационных систем и объектов критической информационной инфраструктуры исключить применение больших фундаментальных моделей искусственного интеллекта с неприемлемым уровнем потенциального ущерба;

- для моделей с высоким уровнем риска предусмотреть дополнительные меры подтверждения безопасности: независимую сертификацию, использование доверенных датасетов, обеспечение технической и криптографической защиты информации по высшим классам;

- для моделей с незначительным уровнем потенциального ущерба допустить использование общедоступных моделей и датасетов и добровольную проверку соответствия требованиям по безопасности.

4.2. Ограничения по применению (суверенные и национальные модели)

В законе [10] определены понятия суверенной и национальной модели искусственного интеллекта, однако в тексте законопроекта они дальнейшего развития не получили. Предусмотрено лишь, что в дальнейшем правовыми актами правительства РФ будет определено, в каких случаях необходимо использовать исключительно национальные или суверенные БФМ искусственного интеллекта. В связи с этим, представляется целесообразным определить поэтапный переход на их использование, начиная с государственных информационных систем и объектов критической информационной инфраструктуры. При определении сроков перехода необходимо учитывать сложность проблемы и степень готовности всей инфраструктуры, прежде всего необходимого аппаратного обеспечения ЦОД и доверенного общесистемного программного обеспечения.

Предлагается создать национальную систему сбора данных и разработать соответствующую нормативно-методическую базу в части порядка сбора и подготовки данных, создать национальный репозиторий доверенных данных для обучения и тестирования систем искусственного интеллекта различного целевого назначения, доверенные бенчмарки в различных предметных областях. Предлагаемые действующими стандартами [17-20] методы оценки качества данных и систем искусственного интеллекта требуют участия квалифицированных экспертов как в области технологии искусственного интеллекта, так и в соответствующих предметных областях. Представляется целесообразным в нормативно-методических документах формализовать и упростить эти процедуры для конкретных применений технологии искусственного интеллекта. При этом должны быть решены проблемы долговременного хранения данных, поддержания их в актуальном состоянии, безопасного санкционированного доступа и др.

Требуется создание национальных стандартов для алгоритмов, интерфейсов, форматов данных и методов оценки качества данных и моделей. Также актуальна проблема совместимости СИИ с существующими корпоративными информационными системами. Требуется разработка отечественного инструментария для разработки моделей ИИ. Само понятие суверенной и национальной модели нежизнеспособно без соответствующего доверенного аппаратного и общесистемного программного обеспечения (операционные системы, BIOS/UEFI, оркестранты, микропрограммное обеспечение и пр.).

В целом, нормативно-методическое регулирование в области разработки и применения БФМ искусственного интеллекта возможно только во взаимной увязке с проблемами разработки и безопасного применения СИИ, включая их общесистемное программно-аппаратное обеспечение. Ошибочное представление о возможности решения постфактум проблем информационной безопасности и комплексной безопасности в целом уже неоднократно приводило к катастрофическим последствиям.

Заметим, что несмотря на значительное смягчение требований к суверенным и национальным моделям в их определении сохранились некоторые неточности. Так, не ясно является ли разработчиком модели потребитель, использующий готовую модель стороннего разработчика с использованием RAG-технологии или в режиме дообучения на этапе эксплуатации. Существенные характеристики модели, определенные правительством РФ, в ст.6 п.3 абз.2) определяются разработчиком и др.

Ст. 6 п.2 абз.3) для суверенной модели предъявляется требование полной воспроизводимости цикла разработки, включая обучение и матрицу исходных коэффициентов. Однако многочисленные публикации [21-23] обосновывают невозможность полной воспроизводимости из-за наличия большого числа факторов случайности в процессе обучения модели помимо исходных коэффициентов. Так, некоторые процессорные операции (например, редукции в сверточных слоях или матричных умножениях) дают слегка разные результаты из-за особенностей параллельных вычислений, в распределенном обучении порядок суммирования градиентов может отличаться из-за задержек сети и планирования задач, что приводит к микроразличиям, некоторые операции фреймворка не детерминированы даже при включенном детерминированном режиме и др. Эти работы показывают, что некоторые источники вариативности (например, фундаментальный недетерминизм параллельных вычислений или нюансы арифметики с плавающей точкой) в принципе сложно полностью исключить. Поэтому в сообществе сейчас вместо требования абсолютной воспроизводимости рекомендуют тщательно документировать все условия эксперимента, запускать несколько итераций и анализировать, как вариации влияют на итоговые выводы.

Определение суверенной и национальной большой фундаментальной модели искусственного интеллекта предусматривает, что ее существенные характеристики, в том числе программное обеспечение и настраиваемые параметры определяются российским разработчиком. При этом наиболее продвинутые отечественные модели в

настоящее время не удовлетворяют этому требованию. Так, GigaChat 3 Ultra от Сбера построена на основе архитектуры DeepSeek V.3, а Yandex GPT 5 Pro инициирована весами Qwen разработки КНР.

4.3. Безопасность

Термин безопасность трактуется в ст.4 (п.9) в самом широком смысле, включая защиту конституционных прав граждан, но в дальнейшем в явном виде не используется. В ст.5 (п.4) упоминается лишь право правительства РФ устанавливать требования по предотвращению рисков, связанных с использованием больших фундаментальных моделей, а в ст.6 предусмотрена проверка соответствия законодательству. Основу законодательства в области защиты информации составляет ФЗ-149 [24], который определяет ее как комплекс правовых, организационных и технических мер, направленных на обеспечение конфиденциальности, целостности и доступности защищаемой информации. Принятые подзаконные акты и нормативно-методические документы уполномоченных федеральных органов исполнительной власти позволяют достаточно адекватно реагировать на современные вызовы и угрозы в этой области. При этом СИИ, как информационные системы, в контексте технической и криптографической защиты информации не обладают какой-либо спецификой [4]. Более того, основные так называемые специфические атаки на системы искусственного интеллекта – отравление, уклонение, реверсивное проектирование и вывода также блокируются или существенно усложняются традиционными средствами технической защиты информации при их грамотном применении на всех стадиях жизненного цикла системы [4].

Однако комплексная безопасность функционирования любой автоматизированной системы и, особенно, системы искусственного интеллекта связана не только с проблемами технической (криптографической) защиты информации. Для многих современных применений информационных технологий на первый план выходят проблемы функциональной безопасности, как способности корректно выполнять заявленные функции в реальных условиях эксплуатации без нанесения неприемлемого ущерба, и проблемы системной безопасности, связанные с корректностью и полнотой постановки функциональных задач [3]. При использовании больших фундаментальных моделей искусственного интеллекта последняя проблема становится особенно актуальной, т.к. качество их функционирования напрямую связано с качеством и полнотой пользовательских запросов – промптов¹, т.е. с уровнем квалификации пользователя. Отсюда вытекает целесообразность разработки учебных программ для специалистов в различных предметных областях и соответствующих квалификационных требований.

Основные проблемы функциональной безопасности [3] при использовании систем искусственного интеллекта связаны с наименее проработанными как в научном, так и в нормативно-методическом плане, обеспечением качества, точности и устойчивости функционирования модели ИИ. На сегодняшний день техническим комитетом ТК-164 разработано более 150 национальных и предварительных стандартов как отраслевых, так и общего характера, в том числе и по указанным направлениям. Однако их полноценное использование возможно только после разработки соответствующих методических рекомендаций. Одновременно, с учетом упомянутого выше предложения о классификации систем искусственного интеллекта в соответствии со степенью потенциального ущерба при некорректном функционировании, необходимо разработать дифференцированные требования к их функциональной безопасности (качеству и устойчивости).

Хорошо известная проблема обеспечения информационной безопасности при эксплуатации информационного ресурса коллективного пользования для систем искусственного интеллекта усугубляется в связи с возможной/неизбежной утечкой корпоративной конфиденциальной информации при формировании запросов. Предлагаемые на сегодняшний день формы и методы обезличивания персональных данных не всегда обеспечивают защиту от полного спектра угроз (получение несанкционированного знания путем санкционированных запросов). Кроме того, необходимо разработать и закрепить в соответствующих стандартах методы обеспечения конфиденциальности иных видов информации ограниченного распространения с одновременным сохранением статистических связей данных для их использования при обучении многопользовательских моделей искусственного интеллекта, например, с использованием гомоморфного шифрования. Целесообразна также разработка соответствующих инструментальных средств для самостоятельной подготовки и разметки данных для обучения (дообучения) владельцем данных, исключающих утечку чувствительной корпоративной информации, в том числе, устойчивых к методам анализа доступной информации OSINT.

4.4. Маркировка контента

Большие изменения претерпел раздел законопроекта о маркировке созданного с использованием больших

¹ По данным Forbs в 46% ответов LLM содержались фактические и в 31% логические ошибки

фундаментальных моделей искусственного интеллекта контента. Теперь маркировка фактически носит добровольный характер и предусматривается только для аудио и визуальной информации. Формат и порядок размещения соответствующего предупреждения определяются соглашением между пользователем и владельцем модели. При этом остается неясным каким образом об этом соглашении узнает потенциальный потребитель этого контента. Очевидно, требуется разработка некоторых стандартов в этой области, с тем чтобы соответствующая метка была доступна как для непосредственного восприятия человеком, так и распознаванию в автоматическом режиме независимо от пользовательской программно-аппаратной среды и технологии обработки контента. В ответственных случаях метка должна также обеспечивать целостность и проверку авторства контента, невозможность ее произвольной модификации. Имеющиеся наработки в этом направлении основаны, как правило, на аппарате хэш-функций. Заметим, что надежность выявления сгенерированного контента, достаточно высокая на начальном этапе эксплуатации соответствующих детекторов, с учетом постоянного совершенствования технологии deepfake, довольно быстро снижается [25]. В целом, маркировка сгенерированного ИИ контента может лишь несколько затруднить, но не исключить распространение фейковой информации. Действительно с задачей создания фейкового контента вполне успешно справляются и локальные модели ИИ без помощи стороннего сервиса ИИ. Кроме того, существующие фильтры больших языковых моделей, препятствующие генерации противоправного контента, достаточно легко обходятся путем с помощью специально подобранной цепочки запросов [26], в том числе с использованием другой модели искусственного интеллекта.

Заключение

Сформулированы предпосылки и факторы, определяющие актуальность нормативного регулирования в области ИИ.

Показана значимость принятого Федерального закона. Он носит рамочный характер и нормативно впервые на федеральном уровне закрепляет базовые понятия и правила игры для разработчиков крупных ИИ-моделей, а именно - больших фундаментальных моделей.

При этом на основе результатов анализа закона, представленных в настоящей статье, можно сформулировать следующие выводы:

Принятый закон [10] не охватывает в полной мере цели и задачи развития искусственного интеллекта в Российской Федерации, определенные Национальной стратегией [7], т.к. сфера его применения ограничена разработкой и применением только одной из технологий ИИ - больших фундаментальных моделей.

В целях реализации положений проекта закона необходимо создание национальной системы сбора, подготовки и использования данных для обучения и тестирования при создании больших фундаментальных моделей искусственного интеллекта.

Необходима разработка подзаконных правовых актов и нормативно-методических документов, определяющих четкие критерии соответствия законодательству в части безопасности применения больших фундаментальных моделей искусственного интеллекта. Такие критерии должны охватывать вопросы технической защиты информации, функциональной и системной безопасности [3,4].

При определении сферы применения национальных и суверенных моделей необходимо учитывать также и состав программно-аппаратной среды их функционирования.

Целесообразно разработать классификацию больших фундаментальных моделей искусственного интеллекта по уровням потенциального ущерба при их применении и соответствующие методические документа для оценки уровня рисков.

В целях повышения качества сгенерированных ответов большой фундаментальной модели искусственного интеллекта целесообразно разработать методические документы, определяющие порядок подготовки промптов и/или корпоративных данных для дообучения модели с целью исключения чувствительной информации.

Необходимо разработать порядок и критерии оценки соответствия больших фундаментальных моделей традиционным российским духовно-нравственным ценностям.

Таким образом, принятый закон об ИИ должен стать нормативно-методической базой, на которой в ближайшие годы будут строиться более конкретные нормы в форме подзаконных и ведомственных документов.

Список литературы

-
1. Стратегия научно-технологического развития Российской Федерации. Утверждена Указом Президента РФ от 28 февраля 2024 г. № 145. – <http://www.kremlin.ru/acts/bank/50358>

2. Путин провел совещание по развитию ИИ в России 10.04.2026 <https://crimea.ria.ru/20260410/putin-provel-soveschanie-po-razvitiyu-ii-v-rossii--glavnye-zayavleniya-prezidenta-1155129276.html>
3. А.А. Зацаринный, В.Е. Гаврилов. Проблемы нормативно-правового и технического регулирования обеспечения информационной безопасности при создании автоматизированных систем военного назначения. Материалы 6-й Международной межведомственной научно-практической конференции РАРАН, Москва, 18 марта 2021 года, Т.2, с.69-75.
4. В.Е. Гаврилов, А.А. Зацаринный. Особенности обеспечения функциональной безопасности автоматизированных систем с применением технологии искусственного интеллекта ССИ Т.34, №3, 2024.
5. А.А. Зацаринный, В.Е. Гаврилов. Некоторые подходы к развитию технического регулирования в области информационной безопасности систем с искусственным интеллектом//ИТ-Стандарт. 2025. № 3 (44). С. 11-18.
6. Федеральный закон РФ (проект) «Об искусственном интеллекте в Российской Федерации и внесении изменений в отдельные законодательные акты» (редакция от 18.03.2026г.) <https://zsrfr.ru/uploads/files/Проект%20ФЗ%20об%20ИИ.pdf>
7. Национальная стратегия развития искусственного интеллекта на период до 2030 года (в редакции Указа Президента Российской Федерации от 15.02.2024 №124), <http://www.kremlin.ru/acts/bank/50326>
8. Федеральный закон РФ (проект) «Об искусственном интеллекте в Российской Федерации и внесении изменений в отдельные законодательные акты Российской Федерации (редакция от 26.05.2026г.) <https://zsrfr.ru/uploads/files/Проект%20ФЗ%20об%20ИИ.pdf>
9. Совет Федерации одобрил Федеральный закон «О поддержке развития технологий искусственного интеллекта в РФ», 17 июля 2026 г. <http://council.gov.ru/events/committees/176242/>
10. Федеральный закон от 26.07.2026 № 243-ФЗ «О поддержке развития технологий искусственного интеллекта в РФ» (принят Госдумой 8.07.2026, одобрен СФ 17.07.2026, утв. Президентом РФ 26.07.2026) <http://publication.pravo.gov.ru/document/0001202607260003?index=1>
11. Advancing United States Leadership in Artificial Intelligence Infrastructure, Executive Order 14141 of January 14, 2025 URL: <https://www.federalregister.gov/documents/2025/01/17/2025-01395/advancing-united-states-leadership-in-artificial-intelligence-infrastructure>
12. Меморандум Министерства войны США об ускорении укрепления доминирующего положения США в области военного ИИ (Перевод профессора И.В. Понкина) <https://vvesti.com/politika/perevod-memoranduma-ministerstva-vojny-ssa-ob-uskorenii-ukreplenia-dominiruusego-polozenia-ssa-v-oblasti-voennogo-ii>
13. ГОСТ Р 71476-2024 (ИСО/МЭК 22989:2022) Искусственный интеллект. Концепции и терминология искусственного интеллекта
14. Касперская: закон об ИИ не поможет России догнать США и Китай <https://dzen.ru/a/ahULTRgISluI6qO>
15. Модельный закон «О критически важных объектах информационно-коммуникационной инфраструктуры», принят постановлением Межпарламентской Ассамблеи государств — участников СНГ от 28 ноября 2014 года №41-14;
16. Regulation (EU) 2024/1689 (2024). Artificial Intelligence Act. Official Journal of the European Union, L 2024/1689;
17. ГОСТ Р 71484.2-2024 (ИСО/МЭК 5259-2-2023) искусственный интеллект. Качество данных для аналитики и машинного обучения. Часть 2. Показатели качества данных;
18. ГОСТ Р 71484.3-2024 (ИСО/МЭК 5259-3-2024) Искусственный интеллект. Качество данных для аналитики и машинного обучения. Часть 3. Требования и рекомендации по управлению качеством данных.
19. ГОСТ Р 59898-2021 Оценка качества систем искусственного интеллекта. Общие положения.
20. 1.11.164-1.161.22 Оценка качества систем искусственного интеллекта. Методы оценки. Первая редакция.
21. H. Semmelrock, S. Kopeinik, D. Theiler, T. Ross-Hellauer, D. Kowald Reproducibility in Machine Learning-Driven Research <https://arxiv.org/abs/2307.10320> Computer Science > Machine Learning Submitted on 19 Jul 2023.
22. S. Kapoor, A. Narayanan, Leakage and the reproducibility crisis in machine-learning-based science. Patterns. A Cell Press journal, Volume 4, Issue 9100804September 08, <https://www.cell.com/patterns/fulltext/HTTPS://S2666-3899%2823%2900159-9?UID=UID%3AA1520953-3278-4F02-BAAA-71332A2E9ED6>
23. A. Desai, M. Abdelhamid, What is reproducibility in artificial intelligence and machine learning research? April 2025 AI Magazine 46(2) DOI:10.1002/aaai.70004
24. Федеральный закон "Об информации, информационных технологиях и о защите информации" от 27.07.2006 N 149-ФЗ;

25. N. Sharma L. Updated. How Accurate are AI Content Detectors and Why are They Important? 29 Nov, 2024. <https://www.analyticsvidhya.com/blog/2024/01/ai-content-detectors/>
26. G. Deng, Y. Liu, Y. Li, K. Wang, Y. Zhang, Z. Li, H Wang, T. Zhang, Y. Liu. MasterKey: Automated Jailbreak Across Multiple Large Language Model Chatbots. Computer Science > Cryptography and Security, Submitted on 16 Jul 2023. <https://arxiv.org/abs/2307.08715>
27. Зацаринный А.А., Иванов К.В. Ключевые проблемы развития технологий искусственного интеллекта в интересах обеспечения военной безопасности государства. // Известия Российской академии ракетных и артиллерийских наук. 2025. № 3 (138). С. 13-19.

References

1. Strategy of Scientific and Technological Development of the Russian Federation. Approved by the Decree of the President of the Russian Federation of February 28, 2024 No 145. – <http://www.kremlin.ru/acts/bank/50358>
2. Putin held a meeting on the development of AI in Russia 10.04.2026 <https://crimea.ria.ru/20260410/putin-provel-soveschanie-po-razvitiyu-ii-v-rossii--glavnye-zayavleniya-prezidenta-1155129276.html>
3. A.A. Zatsarinny, V.E. Gavrilo. Problems of legal and technical regulation of information security in the creation of automated systems for military purposes. Proceedings of the 6th International Interdepartmental Scientific and Practical Conference of the Russian Academy of Sciences, Moscow, March 18, 2021, Vol. 2, pp. 69-75.
4. V.E. Gavrilo, A.A. Zatsarinny. Features of Ensuring the Functional Safety of Automated Systems Using Artificial Intelligence Technology SSI T.34, No3, 2024.
5. A.A. Zatsarinny, V.E. Gavrilo. Some Approaches to the Development of Technical Regulation in the Field of Information Security of Systems with Artificial Intelligence. 2025. № 3 (44). Pp. 11-18.
6. Federal Law of the Russian Federation (draft) "On Artificial Intelligence in the Russian Federation and Amendments to Certain Legislative Acts" (as amended on 18.03.2026) <https://zsr.ru/uploads/files/Проект%20Ф3%20о6%20ИИ.pdf>
7. The National Strategy for the Development of Artificial Intelligence for the Period up to 2030 (as amended by the Decree of the President of the Russian Federation dated 15.02.2024 No124), <http://www.kremlin.ru/acts/bank/50326>
8. Federal Law of the Russian Federation (draft) "On Artificial Intelligence in the Russian Federation and Amendments to Certain Legislative Acts of the Russian Federation" (version of 26.05.2026) <https://zsr.ru/uploads/files/Проект%20Ф3%20о6%20ИИ.pdf>
9. The Federation Council approved the Federal Law "On Support for the Development of Artificial Intelligence Technologies in the Russian Federation", July 17, 2026 <http://council.gov.ru/events/committees/176242/>
10. Federal Law of 26.07.2026 No 243-FZ "On Support for the Development of Artificial Intelligence Technologies in the Russian Federation" (adopted by the State Duma on 8.07.2026, approved by the Federation Council on 17.07.2026, approved by the Federal Assembly of the Russian Federation). President of the Russian Federation 26.07.2026) <http://publication.pravo.gov.ru/document/0001202607260003?index=1>
11. Advancing United States Leadership in Artificial Intelligence Infrastructure, Executive Order 14141 of January 14, 2025 URL: <https://www.federalregister.gov/documents/2025/01/17/2025-01395/advancing-united-states-leadership-in-artificial-intelligence-infrastructure>
12. Memorandum of the U.S. Department of War on Accelerating the Strengthening of the U.S. Dominant Position in the Field of Military AI (Translated by Professor I.V. Ponkin) <https://vvesti.com/politika/perevod-memoranduma-ministerstva-vojny-ssa-ob-uskorenii-ukreplenia-dominiruusego-polozenia-ssa-v-oblasti-voennogo-ii>
13. GOST R 71476-2024 (ISO/IEC 22989:2022) Artificial intelligence. Concepts and terminology of artificial intelligence.
14. Kaspersky: AI law will not help Russia catch up with the United States and China <https://dzen.ru/a/ahULTRgISluI6qO>
15. Model Law "On Critical Information and Communication Infrastructure Facilities", adopted by the Resolution of the Interparliamentary Assembly of the CIS Member Nations of November 28, 2014 No41-14;
16. Regulation (EU) 2024/1689 (2024). Artificial Intelligence Act. Official Journal of the European Union, L 2024/1689;
17. GOST R 71484.2-2024 (ISO/IEC 5259-2-2023) Artificial intelligence. Data quality for analytics and machine learning. Part 2. Data quality indicators;
18. GOST R 71484.3-2024 (ISO/IEC 5259-3-2024) Artificial intelligence. Data quality for analytics and machine learning. Part 3. Data quality management requirements and best practices.

19. GOST R 59898-2021 Quality Assessment of Artificial Intelligence Systems. General provisions.
20. 1.11.164-1.161.22 Quality assessment of artificial intelligence systems. Assessment methods. First edition.
21. H. Semmelrock, S. Kopeinik, D. Theiler, T. Ross-Hellauer, D. Kowald Reproducibility in Machine Learning-Driven Research <https://arxiv.org/abs/2307.10320> Computer Science > Machine Learning Submitted on 19 Jul 2023.
22. S. Kapoor, A. Narayanan, Leakage and the reproducibility crisis in machine-learning-based science. *Patterns. A Cell Press journal*, Volume 4, Issue 9100804 September 08, <https://www.cell.com/patterns/fulltext/HTTPS://S2666-3899%2823%2900159-9?UUID=UUID%3AA1520953-3278-4F02-BAAA-71332A2E9ED6>
23. A. Desai, M. Abdelhamid, What is reproducibility in artificial intelligence and machine learning research? April 2025 *AI Magazine* 46(2) DOI:10.1002/aaai.70004
24. Federal Law "On Information, Information Technologies and Information Protection" dated 27.07.2006 No. 149-FZ;
25. N. Sharma L.Updated. How Accurate are AI Content Detectors and Why are They Important? 29 Nov, 2024. <https://www.analyticsvidhya.com/blog/2024/01/ai-content-detectors/>
26. G. Deng, Y. Liu, Y. Li, K. Wang, Y. Zhang, Z. Li, H Wang, T. Zhang, Y. Liu. MasterKey: Automated Jailbreak Across Multiple Large Language Model Chatbots. *Computer Science > Cryptography and Security*, Submitted on 16 Jul 2023. <https://arxiv.org/abs/2307.08715>
27. Zatsarinny A.A., Ivanov K.V. Key Problems of the Development of Artificial Intelligence Technologies in the Interests of Ensuring the Military Security of the State. *Proceedings of the Russian Academy of Rocket and Artillery Sciences*. 2025. № 3 (138). Pp. 13-19.