ПРАВОВЫЕ ПРОБЛЕМЫ ИДЕНТИФИКАЦИИ ДЕСТРУКТИВНОЙ ИНФОРМАЦИИ В СЕТИ «ИНТЕРНЕТ»[1]

Чудесенко А.В.[2]

Ключевые слова: *идентификация, системы искусственного интеллекта, информационная безопасность, дипфейки, деструктивная информация, маркировка контента, синтетический контент, идентификационные маркеры, цифровые платформы.*

Аннотация

Цель: *выявить и систематизировать актуальные правовые проблемы, связанные с идентификацией деструктивного контента в глобальной сети «Интернет», а также исследовать системные коллизии, возникающие при квалификации информационных материалов в качестве деструктивного контента и установлении факта их системного деструктивного воздействия, с учетом положений современного российского и зарубежного законодательства, судебной практики и доктринальных источников.*

Методы исследования: *сравнительно правовой метод и методы анализа и синтеза в процессе генезиса современных информационных технологий, в том числе и технологий искусственного интеллекта (ИИ).*

Результаты исследования: *Обоснована необходимость критического осмысления таких категорий, как «идентификация деструктивной информации», «деструктивное информационное воздействие», «дипфейки». Отдельно проанализированы проблемы, связанные с использованием автоматизированных систем мониторинга (ИИ) для идентификации деструктивного контента, включая вопросы их ошибок и ответственности. Выявлено, что алгоритмы машинного обучения, несмотря на высокую скорость обработки данных, демонстрируют существенные ограничения в разграничении маркеров определения деструктивной информации от допустимой для распространения. Анализ показывает, что традиционные подходы к идентификации часто не учитывают контекстуальные и культурные особенности распространения контента, что важно для юридической квалификации категории деструктивной информации. Установлено, что некорректная классификация и отсутствие идентификационных маркеров может нарушать права граждан, а делегирование ответственности разработчикам или операторам систем остается законодательно неурегулированным, что создает серьезные вызовы для правоприменительной практики.*

Новизна: *установлена необходимость разработки механизмов совершенствования российского законодательства и правоприменительной практики в условиях развития технологий ИИ, применяемых для идентификации деструктивной информации. Предполагается алгоритм использования технологий ИИ как дополнительных технических инструментов идентификации деструктивной информации, распространяемой в различных источниках.*

DOI: 10.24412/1994-1404-2026-3-00-20

Введение

Цифровая трансформация общественных отношений привела не только к расширению возможностей создания и распространения информации, но и к качественному изменению самой структуры информационного воздействия. Пользователь сети «Интернет» взаимодействует уже не с относительно обозримым массивом сообщений, подготовленных редакциями средств массовой информации или индивидуальными авторами, а с непрерывно обновляемой средой, содержание которой формируется одновременно людьми, алгоритмическими рекомендательными системами и генеративными моделями. В этих условиях

[1] Статья написана в рамках Государственного задания FMUZ-2024-0035 «Обеспечение цифрового суверенитета и информационной безопасности правовыми средствами».

[2] ***Чудесенко Антон Вадимович****, аспирант Института государства и права Российской академии наук (ИГП РАН), младший научный сотрудник сектора информационного права и международной информационной безопасности ИГП РАН, г. Москва, Российская Федерация.*

E-mail: a.chudesenko@mail.ru

дефицит времени и навыков для проверки источника, контекста и достоверности контента становится самостоятельным фактором уязвимости личности. Поэтому проблема деструктивного контента не исчерпывается наличием отдельного запрещенного материала: она связана также со скоростью его воспроизводства, масштабом адресного распространения и способностью цифровой инфраструктуры придавать ему видимость достоверного сообщения.

В соответствии с положениями Стратегии национальной безопасности Российской Федерации (Указ Президента РФ от 02.07.2021 № 400)[3] указанные процессы развиваются в условиях, когда обеспечение информационной безопасности, укрепление государственного суверенитета и достижение технологической независимости отнесены к числу приоритетов государственной политики. Стратегия национальной безопасности Российской Федерации указывает на распространение заведомо недостоверной информации, пропаганду противоправного поведения и использование информационно-коммуникационных технологий для деструктивного воздействия на личность, общество и государство. В свою очередь, Стратегия научно-технологического развития Российской Федерации[4] связывает ответ на большие вызовы с освоением передовых цифровых технологий, обработкой больших данных и созданием отечественных технологических решений. Следовательно, применение искусственного интеллекта (ИИ) в информационной сфере имеет двойственную правовую природу: те же инструменты, которые увеличивают производительность создания контента и позволяют имитировать реальные объекты, могут использоваться для автоматизированного выявления противоправных материалов, проверки их происхождения и снижения нагрузки на ресурсы, осуществляющие идентификацию и модерацию деструктивной информации [1, с. 85].

В настоящее время обеспечение развития научно-технологической политики непосредственно связанно с обеспечением государственного, информационного и технологического суверенитета. Т.А. Полякова и Н.А. Троян справедливо рассматривают научно-технологическую политику не как изолированное направление поддержки инноваций, а как механизм формирования собственной инфраструктуры развития, способной обеспечить национальные интересы в условиях цифровой трансформации [2, с. 195]. Такой подход особенно значим применительно к системам анализа контента: зависимость от иностранных моделей, закрытых наборов данных и непрозрачных критериев модерации способна воспроизводить внешние ценностные и правовые установки внутри национального информационного пространства.

Одновременно возрастает криминогенное использование цифрового контента. Т.А. Полякова и А.А. Смирнов отмечают ускоренное изменение форм киберпреступности, ее транснациональный характер и необходимость межотраслевого развития правовых средств противодействия [3, с. 238].

Важно отметить, что развитие генеративных технологий производства синтетического контента существенно снижает стоимость производства правдоподобных аудиовизуальных подделок и любых других видов информации. Опубликованные в августе 2026 г. результаты отраслевого опроса показывают, что ИИ при подготовке контента используют 73% опрошенных компаний и 45% опрошенных блогеров[5]. Эти данные характеризуют распространенность инструмента среди участников конкретного исследования, но не долю синтетических материалов в сети «Интернет». Тем не менее они подтверждают, что машинная генерация перестала быть исключительной практикой и требует устойчивого режима правового регулирования.

В корпоративном секторе эта проблема дополняется рисками компрометации деловой переписки, подмены управленческих распоряжений и воздействия на работников. А.А. Карцхия и Г.И. Макаренко связывают современную парадигму кибербезопасности с соединением признаков организованной и экономической преступности, совершаемой с использованием цифровой техники, программного обеспечения и ИИ-агентов [4, с. 151].

Прежде чем перейти к рассмотрению возможностей применения систем искусственного интеллекта в решение задачи идентификации и борьбы с распространение деструктивной информации следует определить правовой режим идентификации и подходы к определению данного понятия.

Идентификация деструктивной информации и маркировка синтетического контента как информационно-правовая задача

При анализе проблем идентификации необходимо обеспечить комплексное понимание дефиниции процесса идентификации. По мнению В.Б. Наумова, идентификация — информационный процесс, направленный на установление субъектного и (или) объектного состава общественных отношений на основе идентификатора или их совокупности в соответствии с содержанием отношений и их правовой природой [5, с. 11]. Таким образом, идентификация представляется как комплексный технологический и юридический процесс, включающий в себя определенные критерии, определяемые

[3] Указ Президента РФ от 02.07.2021 № 400 «О Стратегии национальной безопасности Российской Федерации» // Официальный интернет-портал правовой информации. URL: http://pravo.gov.ru, 11.10.2025.

[4] Указ Президента РФ от 28.02.2024 № 145 «О Стратегии научно-технологического развития Российской Федерации» // Собрание законодательства РФ. 2024. № 10. Ст. 1373.

[5] Игнатьев Д. Более 70% компаний и почти половина блогеров используют ИИ для создания контента // Ведомости. 11.08.2026. URL: https://www.vedomosti.ru/media/articles/2026/08/11/1220163-bolee-70-kompanii-i-pochti-polovina-blogerov-ispolzuyut-ii (дата обращения: 31.08.2026).

идентификатором. В вопросе определения деструктивности информации идентификационные маркеры могут определяться по критериям достоверности информации, наличия в ней элементов, направленных на нанесения ущерба национальным интересам и иным критериям определенных концептуальными положениями о информационно-психологической безопасности и защиты от деструктивного информационного воздействия, определенных как цели реализации национальной политики обеспечения информационного суверенитета Российской Федерации.

Необходимо также отличать объект идентификации от предмета оценки. Объектом идентификации является конкретный информационный материал, его источник, автор, распространитель, изображенное лицо, использованная модель либо совокупность технических признаков происхождения. Состояние защищенности личности, общества и государства не может быть «идентифицировано» как единица контента; оно является предметом системной оценки на основе показателей риска, распространенности угроз, устойчивости инфраструктуры и эффективности правоприменения. Такое разграничение имеет методологическое значение: автоматизированная система способна присвоить материалу вероятностный класс, однако вывод об уровне информационно-психологической безопасности требует более широкого междисциплинарного анализа.

Практическая сложность идентификации определяется масштабом информационного оборота. По сообщению представителя Роскомнадзора, опубликованному «Ведомостями», в 2025 г. доступ был ограничен к 1,289 млн материалов против 810,5 тыс. в 2024 г.[6]; мониторинг охватывал крупнейшие социальные сети, видеосервисы и мессенджеры Эта динамика не доказывает пропорционального роста объема противоправной информации, поскольку на показатель влияют расширение мониторинга, изменение перечней запрещенного контента и совершенствование средств обнаружения. Однако она демонстрирует, что проверка каждого материала исключительно вручную организационно невозможна. В результате ИИ приобретает значение инфраструктурного средства осуществления отбора и выделения критически важных маркеров деструктивного контента, а не автономного субъекта осуществления юридической квалификации.

Особенно сложной задачей при идентификации деструктивного и синтетического контента является обнаружение дипфейков. Поэтому перспективной представляется не ставка на один универсальный детектор, а сочетание нескольких источников доказательственной информации: анализа физических и биометрических несоответствий, проверки аудио-характеристик, сопоставления с известными образцами существующего контента, исследования метаданных и цепочки происхождения, а также оценки контекста публикации. В.Б. Наумов обоснованно связывает противодействие дипфейкам с развитием института идентификации, обеспечением достоверности информации и доверия к цифровым коммуникациям, предлагая несколько информационно-правовых моделей и специальные изменения информационного законодательства [6, с. 143].

Принятие Федерального закона № 243-ФЗ стало первым системным шагом федерального законодателя к регулированию разработки, внедрения и применения больших фундаментальных моделей ИИ. Закон определяет ИИ, большую фундаментальную модель, состав данных, обучение модели и ее разработчика, затрагивает вопросы и идентификации, и маркировки контента, созданного с применением технологий ИИ. Статья 9 указанного закона носит наименование «Маркировка информационного материала, созданного с применением больших фундаментальных моделей искусственного интеллекта». Сам факт появления этой статьи имеет принципиальное значение: законодатель признал информационное предупреждение самостоятельным средством обеспечения прозрачности происхождения контента. Вместе с тем содержание нормы не позволяет утверждать, что в России уже введена универсальная обязательная маркировка ИИ-контента. Часть 1 статьи 9 предусматривает, что лицу, применяющему большую фундаментальную модель для создания материала в аудио- и (или) визуальной форме, обеспечивается возможность размещения предупреждения. Согласно части 2 формат, содержание и порядок его размещения устанавливаются соглашением между пользователем и лицом, предоставляющим возможность применения модели. Часть 3 обязывает владельца крупного сайта, информационной системы или программы, отвечающих указанным в законе критериям аудитории и функционального назначения, обеспечить пользователям возможность размещения такого предупреждения[7]. Таким образом, статья 9 закрепляет обязанность создать техническую возможность, но не прямую обязанность автора воспользоваться ею. Закон также не запрещает намеренное удаление технического маркера и не устанавливает специального состава ответственности за распространение немаркированного синтетического материала. Отсылочная статья 11 говорит об ответственности в соответствии с законодательством Российской Федерации, однако сама по себе не образует достаточного основания для санкции, если конкретная нарушенная обязанность и состав правонарушения не закреплены в иных нормах.

Важно обратить внимание и на срок вступления в силу данных законодательных положений. Федеральный закон № 243-ФЗ в основной части вступает в силу 1 сентября 2026 г., тогда как статьи 8—10, включая

[6] РКН в 2025 году удалил на 59% больше запрещенных материалов // Коммерсант. 19.01.2026. URL: https://www.kommersant.ru/doc/8361576 (дата обращения: 31.08.2026).

[7] Федеральный закон от 26.07.2026 № 243-ФЗ «О поддержке развития технологий искусственного интеллекта в Российской Федерации», ч. 1.2 ст. 13.

регулирование маркировки, начинают действовать с 1 марта 2027 г. Переходный период представляется возможным использовать для подготовки согласованного регулирования, поскольку предоставление платформами несовместимых между собой предупреждений не позволит обеспечить процедуру устойчивой идентификации контента.

Искусственный интеллект в идентификации и модерации деструктивной информации

Массовый информационный оборот объективно требует перехода от последовательной ручной проверки к гибридной модели модерации. Использование систем искусственного интеллекта потенциально снижает трудозатраты человека на просмотр заведомо огромного массива информации и позволяет направить квалифицированных специалистов на сложные случаи регулирования. В этом смысле автоматизация является не отказом от правовой оценки, а способом сохранить ее возможность при объеме данных, который превышает человеческие ресурсы.

На практике российская система мониторинга уже развивается по пути функционального разделения анализа текста, изображений, видео и прогнозной аналитики с использованием технологий автоматизации. В открытых источниках системы «Мир», «Окулус» и «Вепрь» описываются как инструменты сбора и текстового анализа, распознавания аудиовизуальных материалов и прогнозного выявления угроз[8]. Независимо от конкретных технических характеристик этих решений дальнейшее развитие должно ориентироваться не только на увеличение количества обработанных единиц контента, но и на проверяемое качество. Учитывая ранее описанные объемы распространения деструктивной информации, «ручная» обработка входящего информационного потока является физически сложной задачей. Поэтому использование систем ИИ может быть представлено как дополнительная вычислительная мощность для выполнения поставленных задач.

Применение систем ИИ способно повысить точность обнаружения деструктивной информации лишь при соблюдении двух условий: детекторы должны непрерывно обновляться и проверяться на новых типах генерации, включая материалы, специально созданные для обхода обнаружения; техническая вероятность не должна автоматически преобразовываться в юридический факт. Такой порядок позволяет согласовать автоматизацию с подходом, обеспечивающим защиту интересов и прав человека в условиях такого стремительного развития технологий. Развиваемое В.Б. Наумовым право на отказ от цифровых технологий направлено на защиту человека как объективно слабой стороны и предполагает возможность выбора в тех правоотношениях, где безальтернативное применение технологии способно существенно затронуть права и свободы [7, с. 29]. Однако указанное право не означает запрета на автоматизированный анализ общедоступного контента или требования проверять весь информационный поток исключительно вручную. В сфере модерации содержанием гарантии должно стать не право исключить материал из машинного сканирования, а право на уведомление о применении автоматизированного решения, получение понятного основания ограничения и рассмотрение возражения человеком, когда мера затрагивает доступ к информации, аккаунт, репутацию или иные законные интересы пользователя. В научной модели безопасного ИИ в обязательном порядке предусматривают право на отказ, доверие, безопасность и постоянное развитие предметного законодательства образуют взаимосвязанные элементы, а не взаимоисключающие требования [8, с. 6]. Следует отметить, что развитие правового регулирования маркировки и автоматизированной модерации деструктивной информации целесообразно осуществлять посредством согласованного пакета изменений.

Представляется возможным внести дополнения в Федеральный закон № 243-ФЗ: в статье 3 следует определить категории «синтетический информационный материал», «существенное преобразование» и «дипфейк» [9, с. 12]. Статью 9 следует дополнить обязательной маркировкой публично распространяемого синтетического аудио- и визуального материала, если его характер способен сформировать у обычного пользователя представление о достоверном отображении реального лица, события или высказывания.

Следующим шагом может быть представлено дополнение положений ФЗ-149. Для владельцев социальных сетей, видеохостингов и иных крупных платформ следует установить обязанности: принимать и сохранять совместимые маркеры происхождения; отображать предупреждение в интерфейсе; не удалять его при технической обработке файла; предоставлять средство сообщения о поддельном изображении или голосе; проводить проверку по обращению изображенного лица; обеспечивать человеческий пересмотр значимого решения; публиковать агрегированную отчетность о выявлении и маркировке. В законе следует разграничить маркировку, ограничение рекомендательного распространения и ограничение доступа как меры разной интенсивности. Наличие синтетического происхождения само по себе может обосновывать предупреждение, но не автоматическую блокировку.

Также следует обратить внимание на процесс совершенствования национального законодательства в части разработки подзаконных актов [10, с. 18]. Правительство Российской Федерации, используя полномочие устанавливать требования по предотвращению рисков применения больших фундаментальных

[8] Количество удаленных РКН запрещенных материалов в 2025 году выросло на 59% // Ведомости. 19.01.2026. URL: https://www.vedomosti.ru/technology/news/2026/01/19/1170082-kolichestvo-udalennih-rkn-zapreschennih (дата обращения: 31.08.2026).

моделей, может определить категории материалов, подлежащих обязательной маркировке, общие исключения и распределение компетенции федеральных органов[9]. Минцифры России возможно наделить полномочием утвердить единые требования к содержанию и визуальному формату предупреждения, составу машиночитаемых сведений, совместимости платформенных интерфейсов и устойчивости технического маркера. Роскомнадзору — утвердить методику контроля исполнения требований, правила проведения контрольных испытаний, показатели качества автоматизированного обнаружения, формы отчетности и порядок взаимодействия с платформами. Технические параметры могут развиваться в национальных стандартах, однако обязательные элементы, влияющие на права и обязанности, должны иметь прямое нормативное основание.

Предлагаемое распределение функций не означает уравнивание национального законодательства и внутренних правил цифровых платформ в смысле их равенства. Внутренние правила являются частным регулятором и должны приводиться в соответствие с императивными требованиями российского права. Вместе с тем законодательство должно учитывать техническую архитектуру платформ и обеспечивать их специфику. Оптимальной является модель, при которой федеральный закон определяет обязанность, субъектов, основания, основные права и последствия; Правительство — риск-категории и общую процедуру; Минцифры — технологический формат; Роскомнадзор — контрольную методику; платформы — конкретную реализацию интерфейса и внутренний порядок рассмотрения обращений в установленных пределах.

Заключение

Проведенный анализ позволяет констатировать системный характер возникающих информационных, технических и правовых, стоящих перед правовой системой РФ в сфере идентификации деструктивного контента и информационного воздействия. Существующая парадигма регулирования, основанная на устаревших подходах, демонстрирует свою недостаточную эффективность перед создаваемыми в огромных масштабах единиц деструктивного контента, наносящих ущерб защищенности общества и государственным интересам [11, с. 229]. Для преодоления возникшей коллизии необходима комплексная модернизация по следующим направлениям.

1) Первоочередным законодательным решением представляется согласованный пакет поправок в Федеральные законы № 243-ФЗ и № 149-ФЗ. В них следует закрепить понятия «синтетического материала» и дипфейка, установить обязательную маркировку «высокорискового» публичного аудио- и визуального контента; определить обязанности крупных платформ сохранять и отображать маркеры, принимать обращения и обеспечивать финальное утверждение решения человеком. Предлагается определить законодательстве обязательные маркеры идентификации деструктивного контента, к числу которых могут быть отнесены: видимое предупреждение, машиночитаемые метаданные/cookie-данные и файлы, устойчивое свидетельство происхождения (источника происхождения контента — человек или ИИ). Такое сочетание повысит качество идентификации даже при удалении одного из элементов.

2) Представляется возможным внесение изменений в акты стратегического планирования и точечные изменения действующих федеральных законов в части расширения понятий «цифровой», «когнитивной зависимости», «медиасуверенитета», «когнитивного суверенитета». Представляется важным закрепить установить обязанности для владельцев крупных информационных, социальных и сетевых платформ по выявлению и маркировке контента, содержащего элементы скрытого психологического воздействия, а также контента, созданного с помощью ИИ-инструментов. Их обязанности могут быть расширены и закреплены в рамках редакции ФЗ-149 и КоАП РФ в части разработки редакции специализированных составов, относящихся к реализации положений ст. 9 ФЗ-244 с определением границ ответственности указанных субъектов.

3) Реализация развития использование технологий искусственного интеллекта для идентификации и модерации информации, обладающей деструктивным воздействием. Помимо этого, необходимо обеспечить развитие собственных LLM-моделей искусственного интеллекта, расположенных на российских серверах и обучаемых с особенностями восприятия контента и информации в рамках российской ментальной среды [12, с. 35]. Комплексное развитие и поддержка собственных информационных систем и систем ИИ позволит повысить защищенность национального суверенитета в условиях вызовов и угроз, а также поддержать и укрепить уровень информационной безопасности и защиты национальных интересов государства. Такой комплексный подход, сочетающий технологическое развитие, законодательное совершенствование и формирование цифровой культуры, позволит обеспечить информационную безопасность личности и общества в условиях развивающейся цифровой эпохи.

[9] Федеральный закон от 26.07.2026 № 243-ФЗ «О поддержке развития технологий искусственного интеллекта в Российской Федерации», п. 4 ч. 2 ст. 5.

Литература

1. Полякова Т.А., Минбалеев А.В., Наумов В.Б. Современные приоритеты развития информационного права: правовое обеспечение государственного суверенитета и информационной безопасности в информационном пространстве России // Государство и право. 2025. № 1. С. 160—173. DOI: 10.31857/S1026945225010148. EDN: HKYGVI.
2. Полякова Т.А., Троян Н.А. Научно-технологическая политика как механизм обеспечения государственного суверенитета в условиях цифровой трансформации // Государство и право. 2026. № 5. С. 193—203. DOI: 10.7868/S2713039826050164.
3. Полякова Т.А., Смирнов А.А. Противодействие киберпреступности в новой цифровой реальности: вызовы и правовые решения // Государство и право. 2026. № 2. С. 236—246. DOI: 10.7868/S2713039826020184.
4. Карцхия А.А., Макаренко Г.И. Кибербезопасность: корпоративно-правовой аспект // Вопросы кибербезопасности. 2026. № 2 (72). С. 149—156. DOI: 10.21681/2311-3456-2026-2-149-156.
5. Наумов В.Б. Институт идентификации в информационном праве : автореф. дис. ... д-ра юрид. наук: 12.00.13 / Наумов Виктор Борисович. М. : ИГП РАН, 2020. С. 11.
6. Наумов В.Б. Борьба с дипфейками: информационно-правовые модели и законодательные решения // Государство и право. 2026. № 4. С. 142—147. DOI: 10.7868/S2713039826040130.
7. Наумов В.Б. Право на отказ от цифровых технологий в сфере искусственного интеллекта // Вестник Университета им. О.Е. Кутафина (МГЮА). 2024. № 10 (122). С. 26—36. DOI: 10.17803/2311-5998.2024.122.10.026-036.
8. Карцхия А.А., Ивлиев Г.П. Современное понимание искусственного интеллекта: правовой аспект // Интеллектуальная собственность. Промышленная собственность. 2025. № 4. С. 4—12.
9. Наумов В.Б. Десять идей для законопроекта о безопасном искусственном интеллекте в России // Юрист. 2026. № 5. С. 10—17. DOI: 10.18572/1812-3929-2026-5-10-17.
10. Карцхия А.А., Макаренко Г.И. Правовые проблемы применения искусственного интеллекта в России // Правовая информатика. 2024. № 1. С. 4-19. DOI: 10.21681/1994-1404-2024-1-4-19.
11. Карцхия А.А., Макаренко Г.И. Актуальные тренды в развитии искусственного интеллекта // Правовая информатика. 2025. № 3. С. 219-231. DOI: 10.24412/1994-1404-2025-3-00-24.
12. Атабеков А.Р. Обеспечение автономности принятия решения искусственный интеллектом для целей публичных правоотношений // Административное и муниципальное право. 2023. № 1. С. 29-36. DOI: 10.7256/2454-0595.2023.1.39893.

SECTION:
CONSTITUTIONAL LAW

LEGAL PROBLEMS OF IDENTIFYING DESTRUCTIVE INFORMATION ON THE INTERNET

***Chudesenko A.V.*[10]**

Keywords: *identification, artificial intelligence systems, information security, deepfakes, destructive information, content labeling, synthetic content, identification markers, digital platforms.*

Abstract

Purpose of the work: *to identify and systematize current legal issues related to the identification of destructive content on the global Internet, as well as to examine systemic conflicts arising in the qualification of information materials as destructive content and in establishing the fact of their systemic destructive impact, taking into account the provisions of modern Russian and foreign legislation, judicial practice, and doctrinal sources.*

Research methods: *comparative legal method and methods of analysis and synthesis in the context of the genesis of modern information technologies, including artificial intelligence technologies.*

[10] ***Anton V. Chudesenko****, Ph.D. student at the Institute of State and Law of the Russian Academy of Sciences (ISL RAS), Junior Researcher at the Sector of Information Technology Law and International Information Security of the ISL RAS, Moscow, Russian Federation.*
E-mail: a.chudesenko@mail.ru

Research results: *The necessity of a critical rethinking of such categories as "identification of destructive information," "destructive information impact," and "deepfakes" has been substantiated. The problems associated with the use of automated monitoring systems (artificial intelligence) for identifying destructive content, including issues of their errors and liability, are analyzed separately. It has been revealed that machine learning algorithms, despite their high data processing speed, demonstrate significant limitations in distinguishing markers for identifying destructive information from content permissible for distribution. The analysis shows that traditional identification approaches often fail to take into account the contextual and cultural features of content dissemination, which is important for the legal qualification of the category of destructive information. It has been established that incorrect classification and the absence of identification markers may violate citizens' rights, while the delegation of responsibility to developers or system operators remains legally unregulated, creating serious challenges for law enforcement practice.*

Novelty: *the necessity of developing mechanisms for improving Russian legislation and law enforcement practice in the context of the development of artificial intelligence technologies used for identifying destructive information is established. An algorithm is proposed for using artificial intelligence technologies as additional technical tools for identifying destructive information disseminated through various sources.*

References

1. Poliakova T.A., Minbaleev A.V., Naumov V.B. Sovremennye prioritety razvitiia informatsionnogo prava: pravovoe obespechenie gosudarstvennogo suvereniteta i informatsionnoi bezopasnosti v informatsionnom prostranstve Rossii // Gosudarstvo i pravo. 2025. No. 1. S. 160—173. DOI: 10.31857/S1026945225010148. EDN: HKYGVI.
2. Poliakova T.A., Troian N.A. Nauchno-tekhnologicheskaia politika kak mekhanizm obespecheniia gosudarstvennogo suvereniteta v usloviiakh tsifrovoi transformatsii // Gosudarstvo i pravo. 2026. No. 5. S. 193—203. DOI: 10.7868/S2713039826050164.
3. Poliakova T.A., Smirnov A.A. Protivodeistvie kiberprestupnosti v novoi tsifrovoi real'nosti: vyzovy i pravovye resheniia // Gosudarstvo i pravo. 2026. No. 2. S. 236—246. DOI: 10.7868/S2713039826020184.
4. Kartskhiia A.A., Makarenko G.I. Kiberbezopasnost': korporativno-pravovoi aspekt // Voprosy kiberbezopasnosti. 2026. No. 2 (72). S. 149—156. DOI: 10.21681/2311-3456-2026-2-149-156.
5. Naumov V.B. Institut identifikatsii v informatsionnom prave : avtoref. dis. ... d-ra iurid. nauk: 12.00.13 / Naumov Viktor Borisovich. M. : IGP RAN, 2020. S. 11.
6. Naumov V.B. Bor'ba s dipfeikami: informatsionno-pravovye modeli i zakonodatel'nye resheniia // Gosudarstvo i pravo. 2026. No. 4. S. 142—147. DOI: 10.7868/S2713039826040130.
7. Naumov V.B. Pravo na otkaz ot tsifrovykh tekhnologii v sfere iskusstvennogo intellekta // Vestnik Universiteta im. O.E. Kutafina (MGIuA). 2024. No. 10 (122). S. 26—36. DOI: 10.17803/2311-5998.2024.122.10.026-036.
8. Kartskhiia A.A., Ivliev G.P. Sovremennoe ponimanie iskusstvennogo intellekta: pravovoi aspekt // Intellektual'naia sobstvennost'. Promyshlennaia sobstvennost'. 2025. No. 4. S. 4—12.
9. Naumov V.B. Desiat' idei dlia zakonoproekta o bezopasnom iskusstvennom intellekte v Rossii // Iurist. 2026. No. 5. S. 10—17. DOI: 10.18572/1812-3929-2026-5-10-17.
10. Kartskhiia A.A., Makarenko G.I. Pravovye problemy primeneniia iskusstvennogo intellekta v Rossii // Pravovaia informatika. 2024. No. 1. S. 4—19. DOI: 10.21681/1994-1404-2024-1-4-19.
11. Kartskhiia A.A., Makarenko G.I. Aktual'nye trendy v razvitii iskusstvennogo intellekta // Pravovaia informatika. 2025. No. 3. S. 219—231. DOI: 10.24412/1994-1404-2025-3-00-24.
12. Atabekov A.R. Obespechenie avtonomnosti priniatiia resheniia iskusstvennyi intellektom dlia tselei publichnykh pravootnoshenii // Administrativnoe i munitsipal'noe pravo. 2023. No. 1. S. 29—36. DOI: 10.7256/2454-0595.2023.1.39893.