Научная статья
https://doi.org/10.24412/2220-2404-2025-8-4
УДК 340

ИСПОЛЬЗОВАНИЕ ТЕХНОЛОГИЙ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ПРОТИВОДЕЙСТВИИ РАСПРОСТРАНЕНИЮ ЭКСТРЕМИСТСКИХ МАТЕРИАЛОВ В СЕТИ «ИНТЕРНЕТ»: СОВРЕМЕННЫЕ ПОДХОДЫ И ПЕРСПЕКТИВЫ

Андреев Н.А.
Краснодарский университет МВД России

Аннотация. *Статья посвящена анализу применения технологий искусственного интеллекта (ИИ) в борьбе с экстремизмом, особое внимание уделено противодействию распространению экстремистских материалов в сети «Интернет». В условиях глобализации и цифровизации экстремизм приобретает новые формы, характеризующиеся анонимностью, масштабируемостью и высокой скоростью распространения, что создает значительные вызовы для правоохранительных органов. В статье рассматриваются основные подходы и методы, использующиеся для мониторинга и фильтрации цифрового контента с применением ИИ, включая машинное обучение и нейросетевые алгоритмы. Особое внимание уделено анализу практических аспектов применения таких технологий, проблемам правового регулирования, а также рискам, связанным с нарушением прав и свобод личности, включая свободу выражения мнений. Также, рассматриваются вопросы международного сотрудничества и необходимость разработки глобальных стандартов для противодействия цифровому экстремизму. Автор подчеркивает важность создания гибридных моделей модерации контента, в которых ИИ будет выполнять роль первичного фильтра, а финальные решения будут приниматься экспертами.*

Ключевые слова: *искусственный интеллект, сеть «Интернет», экстремистский контент, экстремизм, нейросетевые модели, кибербезопасность, цифровой суверенитет, преступность.*

Финансирование: *инициативная работа.*

Original paper

THE USE OF ARTIFICIAL INTELLIGENCE TECHNOLOGIES IN COUNTERING THE SPREAD OF EXTREMIST MATERIALS ON THE INTERNET: MODERN APPROACHES AND PROSPECTS

Nikolay A. Andreev
Krasnodar University of the Ministry of Internal Affairs of Russia

Abstract. *The article is devoted to the analysis of the use of artificial intelligence (AI) technologies in the fight against extremism, with special attention paid to countering the spread of extremist materials on the Internet. In the context of globalization and digitalization, extremism is taking on new forms characterized by anonymity, scalability and high spread rate, which creates significant challenges for law enforcement agencies. The article discusses the main approaches and methods used for monitoring and filtering digital content using AI, including machine learning and neural network algorithms. Special attention is paid to the analysis of practical aspects of the use of such technologies, the problems of legal regulation, as well as the risks associated with the violation of individual rights and freedoms, including freedom of expression. The issues of international cooperation and the need to develop global standards to counter digital extremism are also discussed. The author emphasizes the importance of creating hybrid content moderation models in which AI will act as the primary filter, and final decisions will be made by experts.*

Keywords: *artificial intelligence, Internet, extremist content, extremism, neural network models, cybersecurity, digital sovereignty, crime.*

Financing:*initiativework.*

Введение.

Стремительная цифровая трансформация общества, интенсивно протекающая в последние десятилетия, оказывает существенное влияние на социальный и правовой ландшафт. Современные цифровые технологии позволяют осуществлять радикализацию, пропаганду насилия и разжигание межнациональной, расовой или религиозной ненависти с беспрецедентной скоростью, анонимностью и охватом целевых аудиторий.

В условиях цифровизации экстремизм трансформируется из локального и преимущественно контактного явления в глобальную и анонимную угрозу. Его проявления отличаются высокой степенью интеграции в повседневные цифровые практики, что затрудняет идентификацию субъектов противоправной деятельности и оперативное реагирование со стороны государственных институтов.

В настоящее время интернет-платформа выступает в качестве одного из наиболее значимых каналов экстремистской пропаганды, обеспечивая возможность реализации идеологического воздействия на широкие социальные группы вне зависимости от их географического положения.

В контексте глобализации, одним из определяющих факторов, формирующих архитектуру цифрового пространства, выступает стремительное развитие информационно-коммуникационных технологий (далее - ИКТ), ключевым элементом которых является сеть Интернет. Последняя трансформировалась в универсальный канал распространения информации, средство коммуникации и важнейший инструмент социального взаимодействия.

Однако наряду с позитивным потенциалом цифровой среды, её ресурсы всё чаще используются для целей, прямо угрожающих общественной и государственной безопасности. В частности, речь идёт о распространении деструктивных идеологий, в том числе экстремистской направленности, радикализации отдельных социальных групп и подрыве основ конституционного строя посредством цифровых технологий. Подобная трансформация характера угроз свидетельствует о необходимости переосмысления подходов к обеспечению криминологической безопасности в информационном пространстве.

Острота проблемы подтверждается и на государственном уровне. Так, в ходе расширенного заседания коллегии Министерства внутренних дел Российской Федерации в марте 2025 года Президент Российской Федерации акцентировал внимание на радикализацию интернет-пространства, подчеркнув, что «национальная и религиозная нетерпимость, агрессивный радикализм представляют угрозу единству и сплочённости многонационального народа России, суверенитету и территориальной целостности страны»[1].

Методология.

В качестве методологической основы исследования был выбран комбинированный подход, сочетающий анализ действующих нормативно-правовых актов, судебной практики и современных тенденций в области применения ИКТ для мониторинга интернет-пространства.

Методы исследования включают:

- контент-анализ: исследование государственных программ и стратегий по борьбе с экстремизмом, включая стратегию и планы Министерства внутренних дел РФ.
- статистический анализ: исследование данных о количестве преступлений экстремистской направленности, совершенных с использованием ИКТ, на основе отчетности правоохранительных органов.
- кейс-метод: анализ примеров внедрения ИИ-систем в систему фильтрации экстремистского контента.

Обсуждение.

В стратегии противодействия экстремизму в Российской Федерации акцентировано внимание на возрастающей активности деструктивных организаций, использующих ИКТ, включая мультимедийные и онлайн-платформы, для вербовки новых участников, организации и координации противоправных действий, пропаганды экстремистской идеологии, а также финансирования соответствующей деятельности.

Особую тревогу вызывает возможность осуществления преступлений экстремистской направленности из-за рубежа, вне юрисдикции национальных правоохранительных органов, что ставит под сомнение эффективность традиционных механизмов уголовно-правового реагирования[2].

О серьёзности проблемы экстремистской радикализации интернет-пространства свидетельствует качественное и количественное изменения в структуре преступности

Несмотря на незначительное снижение в Российской Федерации в 2023 году, преступлений экстремистской направленности - до 1340 фактов, что на 14,4 % ниже показателей 2022 года, отмечается устойчивая тенденция к росту преступлений, совершаемых с использованием сети Интернет. Так, 830 из 1340 преступлений экстремистской направленности (62 %) в 2023 году были совершены с применением ИКТ, что на 15,6% превышает показатели предыдущего года[3].

В 2024 году количество зарегистрированных преступлений экстремистской направленности возросло до 1719 (+28,3 %), из которых 962 совершены с использованием ИКТ, что составило прирост на 15,9% по сравнению с 2023 годом[4].

В Концепции противодействия преступлениям, совершаемым с использованием ИКТ, подчёркивается, что рост таких деяний обусловлен не только совершенствованием технологий, но и высоким уровнем виктимности потенциальных потерпевших. Среди детерминант указываются низкий уровень правовой информированности граждан, использование новых криминальных моделей поведения и наличие технологических условий для обеспечения анонимности субъектов преступных посягательств[4].

Осознание масштабов угрозы информационного экстремизма предопределило необходимость формирования устойчивой системы реагирования как на уровне национальной государственной политики, так и в рамках деятельности транснациональных цифровых корпораций. Так, крупнейшие интернет-компании, включая Meta (Facebook), Google и YouTube, внедряют автоматизированные системы фильтрации и удаления противоправного контента, основанные на принципах машинного обучения [5, с. 45]. Согласно статистическим данным, опубликованным международным порталом Statista, ежегодно в социальной сети Facebook удаляется от 20 до 100 миллионов единиц контента, содержащего элементы языка вражды и экстремистской риторики[6].

Однако эффективность указанных механизмов в контексте национальной юрисдикции подвержена сомнению. Так, в Российской Федерации в 2022 году деятельность компании Meta была признана экстремистской в связи с многочисленными случаями распространения призывов к насилию в отношении российских граждан и военнослужащих, а также распространения радикальных установок, угрожающих общественному согласию и национальной безопасности [7].

Однако нельзя признать решенной проблему информационного экстремизма в интернет-пространстве в самой России. По информации Роскомнадзора, в 2024 году, совместно с профильными ведомствами, было удалено либо заблокировано свыше 228 тыс. материалов, сайтов и их отдельных страниц с запрещенной законом информацией. Всего за 2024 год - почти 800 тыс., что на 19% больше, чем в прошлом году[8]. Такая динамика указывает на устойчивую цифровую активность представителей экстремистских сообществ и требует формирования эффективных алгоритмов предупреждения и нейтрализации угроз в рамках единой системы криминологической безопасности».

Анализ статистических данных уголовного судопроизводства за последние пять лет свидетельствует о существенном росте числа лиц, привлечённых к уголовной ответственности по статьям, предусматривающим наказание за преступления экстремистской направленности (в частности, ст. 280, 282, 282^1, 282^2 УК РФ)[9].

По мнению В.В. Баранова, «данная тенденция демонстрирует двоякий характер: с одной стороны, она может рассматриваться как результат повышения эффективности оперативно-розыскной деятельности правоохранительных органов, с другой - как индикатор реального роста экстремистской активности в цифровом пространстве на фоне усиления геополитической напряженности и увеличения количества резонансных экстремистских акций»[10, с. 115].

Принимая во внимание наличие значительного латентного массива преступлений, совершаемых в цифровой среде, становится очевидным, что традиционные методы противодействия экстремизму утрачивают свою эффективность. Это обусловливает необходимость внедрения инновационных подходов, ориентированных на прогнозирование, выявление и пресечение экстремистской активности в режиме реального времени.

Указанные статистические данные иллюстрируют устойчивую и масштабную активность деструктивных элементов в цифровом пространстве, что свидетельствует о нарастающем характере экстремистской угрозы в сети Интернет. Несмотря на предпринимаемые государствами и частными цифровыми платформами меры, экстремистский контент сохраняет способность к широкому распространению, быстро адаптируясь к новым техническим условиям и обходя действующие системы модерации.

Особую обеспокоенность вызывает тот факт, что современные формы цифрового экстремизма все чаще носят скрытый, завуалированный характер. Радикальные идеологи используют легальные на первый взгляд дискурсы (националистические лозунги, религиозные проповеди, политические заявления), которые, будучи помещёнными в определённый контекст, приобретают откровенно деструктивное содержание. Подобная стратегия значительно осложняет процесс идентификации противоправного контента и требует постоянного совершенствования как алгоритмических, так и экспертных механизмов мониторинга.

Дополнительный вызов представляет собой активное использование децентрализованных интернет-платформ, peer-to-peer сетей, шифруемых мессенджеров и форумов в даркнете, где уровень контроля со стороны государства или частных компаний сведен к минимуму[11]. Эти среды становятся резервуаром формирования экстремистской идеологии, подготовки террористических актов и координации противоправных действий, в том числе с международным участием.

Использование криптовалют и анонимных платёжных систем позволяет экстремистским группам избегать финансовой блокировки и правового преследования, что делает необходимым развитие новых инструментов отслеживания подобных операций с привлечением специалистов в области цифровой криминалистики и анализа блокчейн-технологий.

В этом контексте, становится очевидной необходимость усиления межгосударственного сотрудничества, направленного на унификацию подходов к определению экстремистского контента, обмену оперативной информацией и синхронизации законодательных норм в сфере кибербезопасности. Также, возрастает роль международных организаций, включая ООН, Совета Европы и ОБСЕ, в разработке глобальных стандартов противодействия цифровому экстремизму, а также этических и правовых основ применения технологий искусственного интеллекта (далее - ИИ) для идентификации и удаления деструктивного контента[12].

В данном контексте, технологии ИИ представляют собой перспективный инструмент повышения эффективности противодействия информационному экстремизму. Под ИИ, в рамках настоящего исследования, следует понимать совокупность самообучающихся алгоритмов и нейросетевых моделей, способных в автоматическом режиме осуществлять семантический, визуальный и аудиоанализ массивов цифровой информации.

Указанные технологии способны не только идентифицировать признаки экстремистского контента по заранее заданным параметрам, но и адаптироваться к новым формам выражения деструктивных идеологий за счёт постоянного обучения на основе эмпирических данных[13, с. 91].

Использование ИИ-систем позволяет существенно повысить оперативность мониторинга интернет-пространства, снизить зависимость от человеческого фактора и минимизировать временные издержки при анализе больших объёмов данных. Внедрение таких решений особенно актуально в условиях высокой скорости генерации и распространения информации, характерной для цифровой среды.

Результаты.

Таким образом, интеграция технологий искусственного интеллекта в систему противодействия информационному экстремизму представляет собой одно из ключевых направлений формирования современной криминологической политики в условиях цифровой трансформации общества. Необходимым условием до-

стижения высоких результатов, в данной сфере, является междисциплинарный подход, сочетающий правовые, технические, социологические и психологические аспекты изучения и нейтрализации экстремистских угроз.

На сегодняшний день актуальными остаются проблемы отсутствия унифицированных стандартов обработки цифрового контента, недостаточной транспарентности алгоритмов и рисков ошибочной классификации. Такие риски предполагают возможность необоснованного ограничения свободы выражения мнений, что требует от государств разработки сбалансированных правовых механизмов, обеспечивающих соблюдение принципа пропорциональности при реализации мер цифровой модерации.

В контексте противодействия экстремизму в сети Интернет, технологии ИИ обладают целым рядом преимуществ, включая способность к оперативной обработке больших объёмов данных, высокой степени точности при классификации контента, адаптации к новым моделям поведения злоумышленников и возможности интеграции с национальными системами мониторинга[14].

Вместе с тем, необходимо признать, что существующие алгоритмы, несмотря на свою эффективность, нередко страдают от «чёрного ящика» принятия решений, что затрудняет процесс юридической оценки правомерности удаления контента и привлечения виновных к ответственности.

Особый интерес представляет опыт внедрения ИИ-систем российскими структурами, такими как Роскомнадзор, МВД России и рядом частных организаций, в числе которых наиболее заметную роль играет компания «Крибрум»[15]. Последняя занимается разработкой аналитических платформ, способных в режиме реального времени выявлять экстремистский контент в социальных сетях и других публичных источниках, классифицируя его по тематическим, поведенческим и лексико-семантическим признакам. На основе многолетней выборки и постоянного дообучения такие системы обеспечивают высокий уровень чувствительности к новым угрозам, адаптируясь к меняющимся паттернам радикальной коммуникации.

Вместе с тем, нельзя не упомянуть о рисках, возникающих при внедрении алгоритмических систем модерации, которые должны заслуживать особого внимания со стороны как разработчиков ИИ, так и правоприменителей.

Одним из ключевых вызовов в этой сфере является обеспечение соблюдения фундаментальных прав человека - в первую очередь, свободы выражения мнений, права на доступ к информации и гарантий неприкосновенности частной жизни.

Ошибочные или произвольные решения ИИ-систем по удалению контента могут привести к дискриминации, цензуре и подрыву легитимности регулирующих механизмов, особенно в контексте политически чувствительных тем. Именно в этом контексте, исследователи поднимают вопрос о так называемом digitalcolonialism- феномене, при котором алгоритмы, разработанные транснациональными корпорациями, слабо адаптированы к культурным и языковым особенностям конкретных стран, но при этом активно используются для регулирования информационного пространства на глобальном уровне[16, с. 66]. Подобная ситуация может привести к навязыванию универсалистских норм, не учитывающих локальные контексты, что, в свою очередь, снижает эффективность борьбы с экстремизмом и одновременно усиливает отчуждение граждан от цифровой инфраструктуры, которую они воспринимают как репрессивную.

Кроме того, техническая реализация ИИ-модерации, часто оказывается закрытой для внешнего анализа - исходные коды и параметры обучения не раскрываются, а решения принимаются на основе «чёрного ящика» (black-boxmodels)[17]. Это порождает дополнительные сложности при оценке законности и обоснованности ограничительных мер, особенно в случаях обжалования удаления контента в суде или перед регулятором.

Таким образом, прозрачность алгоритмов, их верифицируемость и подотчётность становятся центральными условиями легитимного применения ИИ в сфере цифровой модерации.

Кроме того, существует и пробельность в нормативно-правовом регулировании: несмотря на активное развитие доктрины цифрового суверенитета и принятие отдельных законодательных актов, в российском праве пока отсутствует целостный нормативный массив, регулирующий использование ИИ в сфере выявления, блокировки и удаления противоправного контента[18, с. 209]. Это препятствует формированию единой правоприменительной практики и создает риск как злоупотреблений, так и правовой неопределённости в действиях операторов цифровых платформ и государственных органов.

Учитывая указанные вызовы, одним из перспективных направлений становится разработка и внедрение гибридных моделей модерации, в которых ИИ-системы могут используются как первичный фильтр, а финальные решения об удалении контента должны приниматься квалифицированными экспертами с учётом правовых и культурных контекстов. Такая модель позволяет сочетать масштабируемость автоматизированной обработки с нюансированным подходом человеческого анализа.

Дополнительно необходимо нормативное закрепление принципов этичного и правомерного использования ИИ в сфере информационной безопасности. Среди них ключевыми являются:

- принцип соразмерности, требующий взвешенности между интересами безопасности и свободой выражения;
- принцип подотчётности, предполагающий наличие механизмов обжалования решений ИИ и ответственности операторов;
- принцип транспарентности, обязывающий раскрывать логику принятия решений и характеристики используемых алгоритмов;

- принцип ненанесения ущерба, предусматривающий предотвращение социальной дискриминации, стигматизации и цензуры в отношении уязвимых групп.

В перспективе, учитывая трансграничный характер угроз и цифровой среды, особенно важно выстраивать правовые рамки на базе международного сотрудничества. Создание глобальных соглашений, кодексов этики и технических стандартов, согласованных между странами и технологическими компаниями, позволит сформировать устойчивую инфраструктуру противодействия экстремизму в Интернете, способную эффективно реагировать на угрозы, сохраняя при этом высокий уровень защиты прав личности.

Заключение.

Подводя итоги полученным результатам анализа практического опыта применения технологий ИИ в борьбе с цифровыми формами экстремизма, представляется возможным констатировать наличие устойчивой тенденции к институционализации алгоритмического контроля в информационной среде. Вместе с тем остаются нерешёнными вопросы нормативной регламентации таких систем, их верификации и независимого аудита, что требует разработки новых подходов к правовому регулированию алгоритмической модерации, основанных на принципах транспарентности, подотчётности и соблюдения баланса между общественной безопасностью и цифровыми правами граждан.

В связи с этим, в заключительной части статьи предлагается сформулировать системные рекомендации по совершенствованию нормативно-правовой базы и институциональных механизмов применения ИИ в целях противодействия распространению экстремистской идеологии в цифровом пространстве. Указанные предложения охватывают как направления модернизации законодательства (включая принятие специальных подзаконных актов и технических регламентов), так и аспекты организационного характера -, например, создание специализированных центров мониторинга на базе межведомственного взаимодействия, включающих представителей органов внутренних дел, регуляторов цифровой среды и независимых экспертных организаций.

Разработка таких рекомендаций должна опираться на принципы соразмерности, правовой определённости, соблюдения процедурных гарантий, а также внедрение института алгоритмического аудита, обеспечивающего контроль за соблюдением прав субъектов цифровых коммуникаций.

Только в условиях формирования сбалансированной нормативной среды возможно достижение устойчивых результатов в противодействии информационному экстремизму с применением современных интеллектуальных технологий.

Конфликт интересов

Не указан.

Рецензия

Все статьи проходят рецензирование в формате double-blind peer review (рецензенту неизвестны имя и должность автора, автору неизвестны имя и должность рецензента). Рецензия может быть предоставлена заинтересованным лицам по запросу.

Conflict of Interest

None declared.

Review

All articles are reviewed in the double-blind peer review format (the reviewer does not know the name and position of the author, the author does not know the name and position of the reviewer). The review can be provided to interested persons upon request.

Список источников:

1. Стенограмма выступления Путина на расширенном заседании коллегии МВД [Электронный ресурс] // Режим доступа: http://www.kremlin.ru/events/president/news/73770(дата обращения: 16.06.2024)

2. Об утверждении Стратегии противодействия экстремизму в Российской Федерации: Указ Президента России от 28 декабря 2024 г. №1124 [Электронный ресурс] // Режим доступа: https://www.garant.ru/products/ipo/prime/doc/411135491/?ysclid=m7od3crvzt81326651(дата обращения: 16.01.2025)

3. Состояние преступности в России за январь – декабрь 2023 года [Электронный ресурс] // Режим доступа: file:///C:/Users/79614/Downloads/Sbornik_dlya_UOS.pdf(дата обращения: 16.06.2024)

4. Краткая характеристика состояния преступности в Российской Федерации за январь - декабрь 2024 года [Электронный ресурс] // Режим доступа: https://мвд.рф/reports/item/60248328/ (дата обращения: 11.05.2025)

5. Касенов А.Д. Фомичев Т.С. Пути противодействия экстремизму в социальных сетях // Человек. Общество. Наука. 2024. Т.5. №4. С. 43-54.

6. Количество удаленных материалов с разжиганием ненависти на Facebook по всему миру с 4-го квартала 2017 года по 4-й квартал 2024 года (в миллионах) [Электронный ресурс] // Режим доступа: https://www.statista.com/statistics/1013804/facebook-hate-speech-content-deletion-quarter/ (дата обращения: 16.06.2025)

7. Суд объяснил, почему признал Meta экстремистской [Электронный ресурс] // Режим доступа: https://smotrim.ru/article/2695320?ysclid=mcqcur5xy5308800643 (дата обращения: 16.06.2025)

8. РКН в 2024 году заблокировал почти 800 тыс. материалов с запрещенной информацией [Электронный ресурс] // Режим доступа: https://tass.ru/obschestvo/22903351?ysclid=mcqd1phlg472147469 (дата обращения: 16.06.2025)

9. Число дел об экстремизме в России за 2024 год выросло на 43% [Электронный ресурс] // Режим доступа: https://ria.ru/20250115/chislo-1993749833.html?ysclid=mcqdbnwph8345944894 (дата обращения: 16.06.2025)

10. Баранов В.В. Использование искусственного интеллекта для выявления и предотвращения распространения экстремистской и террористической пропаганды в Интернете // Труды Академии управления МВД России. 2024. № 4 (72). С. 113–128.

11. Киберпреступность и даркнет: загадочный мир, полный опасностей [Электронный ресурс] // Режим доступа: https://dzen.ru/a/ZQLmvYNbJ0OQsER2?ysclid=mcqe6h3xr3913853750 (дата обращения: 21.06.2025)

12. Что не так с Глобальным цифровым договором [Электронный ресурс] // Режим доступа: https://russiancouncil.ru/analytics-and-comments/analytics/chto-ne-tak-s-globalnym-tsifrovym-dogovorom/?ysclid=mcqe9xudhm843249030 (дата обращения: 21.06.2025)

13. Остроух А.В. Введение в искусственный интеллект: монография. - Красноярск: Научно-инновационный центр, 2020. – 250 с.

14. Экосистема онлайн-экстремизма [Электронный ресурс] // Режим доступа: https://katehon.com/ru/article/ekosistema-onlayn-ekstremizma?ysclid=mcqdror59r760395335 (дата обращения: 23.06.2025)

15. Компания «Крибрум» успешно завершила проект по разработке нового программного продукта «Крибрум. OSINT» [Электронный ресурс] // Режим доступа: https://kribrum.ru/company/ (дата обращения: 23.06.2025)

16. Курочкин А.В., Морозова С.С. Цифровая колониализация как угроза национальной безопасности // ПОЛИТЭКС. 2024. №1. С. 64-72.

17. Проблема «черного ящика» в искусственном интеллекте: можем ли мы доверять тому, что не понимаем? [Электронный ресурс] // Режим доступа: https://dzen.ru/a/ZzsqmavEOV97fh7H?ysclid=mcqeodioqa28219158 (дата обращения: 23.06.2025)

18. Никонов В.А., Воронов А.С., Сажина В.А., Володенков С.В., Рыбакова М.В. Цифровой суверенитет современного государства: содержание и структурные компоненты (по материалам экспертного исследования) // Вестник Томского государственного университета. Философия. Социология. Политология. 2021. № 60. С.206–216.

References:

1. Transcript of Putin's speech at an expanded meeting of the Board of the Ministry of Internal Affairs [Electronic resource] // Access mode: http://www.kremlin.ru/events/president/news/73770 (date of request: 06/16/2024)

2. On the approval of the Strategy for Countering Extremism in the Russian Federation: Decree of the President of Russia dated December 28, 2024 No. 1124 [Electronic resource] // Access mode: https://www.garant.ru/products/ipo/prime/doc/411135491 /?ysclid=m7od3crvzt81326651(accessed: 01/16/2025)

3. The state of crime in Russia in January – December 2023 [Electronic resource] // Access mode: file:///C:/Users/79614/Downloads/Sbornik_dlya_UOS.pdf(accessed: 06/16/2024)

4. Brief description of the state of crime in the Russian Federation in January - December 2024 [Electronic resource] // Access mode: https://мвд.рф/reports/item/60248328 / (date of request: 05/11/2025)

5. Kasenov A.D. Fomichev T.S. Ways of countering extremism in social networks. Society. Science. 2024. Vol. 5. No. 4. pp. 43-54.

6. The number of deleted hate speech materials on Facebook worldwide from the 4th quarter of 2017 to the 4th quarter of 2024 (in millions) [Electronic resource] // Access mode: https://www.statista.com/statistics/1013804/facebook-hate-speech-content-deletion-quarter / (date of appeal: 06/16/2025)

7. The court explained why it recognized Meta as extremist [Electronic resource] // Access mode: https://smotrim.ru/article/2695320 ?ysclid=mcqcur5xy5308800643 (accessed: 06/16/2025)

8. In 2024, the RCN blocked almost 800 thousand materials with prohibited information [Electronic resource] // Access mode: https://tass.ru/obschestvo/22903351 ?ysclid=mcqd1phlg472147469 (accessed: 06/16/2025)

9. The number of cases of extremism in Russia increased by 43% in 2024 [Electronic resource] // Access mode: https://ria.ru/20250115/chislo-1993749833.html ?ysclid=mcqdbnwph8345944894 (accessed: 06/16/2025)

10. Baranov V.V. The use of artificial intelligence to identify and prevent the spread of extremist and terrorist propaganda on the Internet // Proceedings of the Academy of Management of the Ministry of Internal Affairs of Russia. 2024. No. 4 (72). pp. 113-128.

11. Cybercrime and the darknet: a mysterious world full of dangers [Electronic resource] // Access mode: https://dzen.ru/a/ZQLmvYNbJ0OQsER2 ?ysclid=mcqe6h3xr3913853750 (accessed: 06/21/2025)

12. What's wrong with the Global Digital Contract [Electronic resource] // Access mode: https://russiancouncil.ru/analytics-and-comments/analytics/chto-ne-tak-s-globalnym-tsifrovym-dogovorom/?ysclid=mcqe9xudhm843249030 (accessed: 06/21/2025)

13. Ostroukh A.V. Introduction to artificial intelligence: monograph. Krasnoyarsk: Scientific and Innovation Center, 2020. 250 p.

14. Ecosystem of online extremism [Electronic resource] // Access mode: https://katehon.com/ru/article/ekosistema-onlayn-ekstremizma ?ysclid=mcqdror59r760395335 (accessed: 06/23/2025)

15. The Kribrum company has successfully completed a project to develop a new Kribrum software product. OSINT" [Electronic resource] // Access mode: https://kribrum.ru/company / (date of access: 06/23/2025)

16. Kurochkin A.V., Morozova S.S. Digital colonization as a threat to national security // POLITEX. 2024. No. 1. pp. 64-72.

17. The problem of the "black box" in artificial intelligence: can we trust what we do not understand? [Electronic resource] // Access mode: https://dzen.ru/a/ZzsqmavEOV97fh7H ?ysclid=mcqeodioqa28219158 (accessed: 06/23/2025)

18. Nikonov V.A., Voronov A.S., Sazhina V.A., Volodenkov S.V., Rybakova M.V. Digital sovereignty of the modern state: content and structural components (based on expert research) // Bulletin of Tomsk State University. Philosophy. Sociology. Politicalscience. 2021. No. 60. pp.206-216.

Информация об авторе:
Андреев Николай Александрович, адъюнкт кафедры уголовного права и криминологии Краснодарского университета МВД России, milena.555@mail.ru
Nikolay A. Andreev, Associate Professor of the Department of Criminal Law and Criminology, Krasnodar University of the Ministry of Internal Affairs of Russia.

Статья поступила в редакцию / The article was submitted 07.07.2025;
Одобрена после рецензирования / Approved after reviewing 23.07.2025;
Принята к публикации / Accepted for publication 20.08.2025.
Автором окончательный вариант рукописи одобрен.