

3. Применение информационных технологий в профилактике экстремизма в молодежной среде : учебно-методическое пособие / П. Е. Суслонов, А. Н. Старостин, В. Л. Назаров [и др.] ; под общ. ред. В. Л. Назарова ; Министерство науки и высшего образования Российской Федерации. Екатеринбург : Изд-во Урал.ун-та, 2021. 140 с.

4. Прокопьева А. В. Экстремизм и его угрозы в отношении молодежи Иркутской области // Экстремизм: новые вызовы региональной безопасности : сборник научных трудов по итогам международной научно-практической конференции (Барнаул, 27 марта 2025 года). Барнаул : Барнаульский юридический институт МВД России, 2025. С. 93–95.

5. Профилактика экстремизма в сети Интернет // Международный институт компьютерных технологий. URL: <https://iict.ru/> (дата обращения: 28.02.2026).

© Прокопьева А. В., 2026

УДК 323.28:343.85

Рябчиков С. А. – доцент кафедры криминалистики Уфимского юридического института МВД России (Россия, г. Уфа)

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В ПРОФИЛАКТИКЕ ЭКСТРЕМИЗМА: ТЕХНОЛОГИЧЕСКИЕ ВОЗМОЖНОСТИ И ПРАВОВЫЕ ОГРАНИЧЕНИЯ

ARTIFICIAL INTELLIGENCE IN EXTREMISM PREVENTION: TECHNOLOGICAL CAPABILITIES AND LEGAL RESTRICTIONS

Аннотация. В статье проводится комплексный анализ применения технологий искусственного интеллекта в профилактике экстремистской деятельности в цифровой среде. Исследуются современные технологические решения для мониторинга, выявления и пресечения распространения экстремистских материалов в сети Интернет, отмечается положительный опыт Российской Федерации. Подробно рассматриваются правовые ограничения, связанные с использованием искусственного интеллекта в деятельности сотрудников правоохранительных органов, включая защиту персональных данных, прозрачности алгоритмов и соблюдения фундаментальных прав человека. На основе сравнительного анализа международных и национальных правовых норм формулируются рекомендации по сбалансированному развитию практических подходов, позволяющих реализовать потенциал искусственного интеллекта по профилактике экстремизма.

Ключевые слова: искусственный интеллект, экстремизм, терроризм, цифровая среда, объяснимый искусственный интеллект, правовое регулирование, система «Окулус», «Крибрум».

Abstract. The article provides a comprehensive analysis of the use of artificial intelligence technologies in the prevention of extremist activity in the digital environment. The article examines modern technological solutions for monitoring, detecting and suppressing the spread of extremist materials on the Internet, and highlights the positive experience of the Russian Federation. The legal restrictions related to the use of artificial intelligence in the activities of law enforcement officers, including the protection of personal data, transparency of algorithms and respect for fundamental human rights, are considered in detail. Based on a comparative analysis of international and national legal norms, recommendations are formulated for the balanced development of practical approaches that make it possible to realize the potential of artificial intelligence to prevent extremism.

Keywords: artificial intelligence, extremism, terrorism, digital environment, explicable artificial intelligence, legal regulation, Oculus system, Kribrum.

Современная цифровая трансформация общества создала беспрецедентные возможности для коммуникации и обмена информацией, но одновременно сформировала благоприятную среду для распространения экстремистской и террористической идеологии [1, с. 419]. Массовый и анонимный характер интернет-пространства способствует разжиганию ненависти, вербовке сторонников радикальных движений и координации противоправной деятельности. Традиционные подходы к противодействию, основанные на ручной модерации контента и точечных мерах реагирования, демонстрируют свою неэффективность перед масштабами и адаптивностью онлайн-кампаний экстремистского характера.

В этих условиях технологии искусственного интеллекта (далее – ИИ) рассматриваются как перспективный инструмент для автоматизации процессов выявления, анализа и блокировки противоправных материалов. Актуальность статьи обусловлена стремительным внедрением ИИ-систем в правоохранительную практику при одновременном нарастании дискуссий об этических границах их применения. Целью исследования являются комплексный анализ технологических возможностей ИИ в профилактике экстремизма и критическое осмысление связанных с этим этико-правовых ограничений, что позволит сформулировать основы сбалансированного подхода к регулированию данной сферы.

Современные системы ИИ обеспечивают возможность автоматического анализа огромных массивов текстовых, аудиовизуальных и мультимедийных данных в режиме, близком к реальному времени. Наиболее пер-

спективными методами являются классификация на основе машинного обучения и сопоставление с базами данных «хешей» известного запрещенного контента. Алгоритмы глубокого обучения, в частности нейронные сети, применяются для бинарной классификации сообщений в социальных сетях на предмет содержания признаков экстремистской деятельности. Ключевое преимущество ИИ заключается в способности обрабатывать объемы информации, недоступные для человеческого восприятия, выявляя сложные закономерности и аномалии в поведенческих паттернах пользователей [2, с. 125].

Системы предиктивной аналитики на основе машинного обучения позволяют не только идентифицировать уже существующие угрозы, но и прогнозировать потенциальные сценарии развития экстремистской активности, осуществляя проактивное предотвращение преступлений. Международный опыт, в том числе инициативы Европейского союза по разработке системы RED-Alert, использующей ИИ для обнаружения террористического контента онлайн, подтверждает глобальный масштаб применения подобных технологий [3].

В Российской Федерации активно развивается и внедряется комплекс специализированных ИИ-систем, направленных на противодействие экстремизму в цифровом пространстве. Система «Окулус» предназначена для автоматического распознавания запрещенной информации в изображениях и видеоконтенте [4]. Технические возможности системы позволяют обрабатывать более 200 000 изображений в сутки, затрачивая на анализ одной единицы контента около 3 секунд, что на порядки превышает производительность ручной проверки. Система обучена распознавать широкий спектр нарушений: экстремистский контент, призывы к массовым беспорядкам, пропаганду наркотических веществ и суицидальный контент.

Аналитическая платформа «Крибрум» используется для мониторинга социальных сетей и публичных источников в режиме реального времени [5]. Она классифицирует контент по тематическим, поведенческим и лексико-семантическим признакам, обеспечивая высокую чувствительность к новым формам радикальной риторики за счет постоянного дообучения на актуальных данных. Проекты МВД России, такие как «Клон» и «Конъюнктура», направлены на потоковый анализ данных и деанонимизацию участников экстремистской деятельности в сети [6].

В соответствии со Стратегией противодействия экстремизму в Российской Федерации до 2030 года, утвержденной Указом Президента Российской Федерации от 28 декабря 2024 г. № 1124 [7], приоритетом является проведение научных исследований для разработки и внедрения технологий ИИ в мониторинговую деятельность правоохранительных органов. Крупные технологические и медиакомпании также интегрируют ИИ в процессы внутреннего контроля. В частности, издательство «Эксмо» использует многоэтапную ИИ-систему для проверки текстов на наличие за-

прещенного контента, что позволяет на ранней стадии выявить потенциальные проблемы еще до публикации.

Применение ИИ в профилактике экстремизма сопряжено с правовыми и этическими нормами, которые необходимо учитывать при формировании регуляторной базы. Одним из наиболее острых является риск алгоритмической предвзятости и дискриминации. Если наборы данных не являются репрезентативными или отражают исторически сложившиеся стереотипы, алгоритмы могут вырабатывать дискриминационные решения в отношении определенных социальных, этнических или религиозных групп. Системы оценки рисков, применяемые в правоохранительной деятельности, способны некорректно идентифицировать в качестве «потенциально опасных» представителей меньшинств, против которых уже имелаась предвзятая статистика. Исследования показывают, что предвзятость может возникать на трех уровнях: предвзятость данных (несбалансированные выборки), алгоритмическая предвзятость (системные ошибки в моделях) и когнитивная предвзятость (неосознанное внесение разработчиками собственных стереотипов) [8, с. 494].

Массовый мониторинг цифровых коммуникаций с помощью ИИ-систем создает системную угрозу праву на частную жизнь, гарантированному как национальными конституциями, так и международными правовыми актами. Технологии глубокого анализа данных, деанонимизации и прогнозного моделирования поведения фактически реализуют модель превентивного надзора, когда подозрения возникают не на основе конкретных действий, а на основе вероятностных прогнозов алгоритма. Подобные практики несут риск превращения цифрового пространства в среду тотального контроля, что противоречит фундаментальным принципам правового государства.

Не менее значимой проблемой является так называемый эффект «черного ящика»: многие современные сложные модели ИИ, особенно глубокого обучения, функционируют таким образом, что даже их создатели не всегда могут детально объяснить, как именно алгоритм пришел к тому или иному выводу. В правовом поле это создает непреодолимое препятствие для реализации принципов состязательности и права на защиту. Если решение о блокировке контента или применении иных мер существенно зависит от рекомендации непрозрачной ИИ-системы, сторона защиты лишается возможности оспорить логику и основания этого решения. Ответом на данный вызов является развитие методов объяснимого ИИ (Explainable AI – XAI), направленных на создание инструментов, делающих внутреннюю работу алгоритмов понятной для юриста, судьи или подозреваемого [9, с. 292].

Применение ИИ в профилактике экстремизма сталкивается также с проблемой размытости самих правовых дефиниций. Понятия «экстремизм» и «разжигание ненависти» нередко зависят от культурного и поли-

тического контекста, что затрудняет формализацию критериев для обучения алгоритмов. Алгоритм, настроенный на выявление экстремизма, может ошибочно классифицировать как противоправную критику власти, сатиру или академическую дискуссию. В итоге возникает фундаментальный конфликт между двумя конституционно защищаемыми ценностями: безопасностью общества и свободой выражения мнений.

Несмотря на высокую производительность, ИИ-системы демонстрируют ограниченную способность к пониманию глубокого контекста, иронии, сарказма и культурных особенностей коммуникации. Это приводит к значительному проценту ложноположительных срабатываний. Например, система в издательстве «Эксмо» по ошибке пометила как нарушителя писателя Дениса Драгунского из-за совпадения начала его фамилии с английским словом «drug» (наркотик) [10]. В результате эффективность систем часто зависит не от самого ИИ, а от выстроенной вокруг него процедуры с обязательной последующей проверкой человеком.

Российская Федерация последовательно формирует собственную регуляторную базу. В 2026 г. на рассмотрении находился проект федерального закона «Об основах государственного регулирования сфер применения технологий искусственного интеллекта в Российской Федерации», вступление в силу которого планируется с 1 сентября 2027 г. [11]. Законопроект предусматривает риск-ориентированный подход к регулированию, введение понятий «суверенных», «национальных» и «доверенных» моделей ИИ, право граждан на досудебное обжалование решений, принятых с использованием ИИ, и ответственность разработчиков и операторов за результаты работы систем. Этот закон призван заложить правовой фундамент для безопасного и этичного развития ИИ-технологий в стране, включая сферу обеспечения безопасности.

На основании проведенного исследования стоит отметить, что применение ИИ в профилактике экстремизма представляет собой мощный, но неоднозначный инструмент, находящийся на пересечении технологического прогресса, права и этики. Технологические возможности ИИ от автоматического мониторинга гигантских массивов данных в реальном времени до прогнозной аналитики открывают качественно новые перспективы для превентивного противодействия радикализации в цифровой среде. Российская практика, в частности системы «Окулус» и «Крибрум», демонстрирует высокую операционную эффективность подобных решений.

Вместе с тем реализация этого потенциала основывается на риске алгоритмической дискриминации, угрозе фундаментальным правам человека, прежде всего праву на приватность и свободу выражения мнений, проблемой «черного ящика», недостаточной прозрачности решений и высоким уровнем ложных срабатываний.

Успешное внедрение ИИ в практику профилактики экстремизма требует сбалансированного подхода, предполагающего правовые рамки с ме-

ханизмами контроля и ответственности, приоритет человеческого контроля над окончательными решениями, развитие технологий объяснимого ИИ (ХАИ), усиление международного сотрудничества в противодействии трансграничному цифровому экстремизму. Только при соблюдении этих условий ИИ сможет стать не инструментом тотального контроля, а соразмерным средством защиты общества при безусловном сохранении верховенства права и ценности человеческой свободы.

Библиографический список

1. Меняйло Д. В., Крупенникова К. К., Меняйло Л. Н. Трансформация экстремизма в условиях цифровизации общества // Право и государство: теория и практика. 2023. № 11 (227). С. 417–419.

2. 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.

3. RED-Alert; the real-time detection and alert system for online terrorist content // Insikt Intelligence. 2018. URL: <https://insiktintelligence.com/red-alert-system-for-online-terrorist-content/> (дата обращения: 02.03.2026).

4. Что такое «Окулус». Как Роскомнадзор использует новую систему слежки // ZakonGuru. 2023. URL: https://dzen.ru/a/Y_-bZ6TA02U68Z3- (дата обращения: 02.03.2026).

5. Сайт компании «Крибрум». URL: <https://kribrum.ru/> (дата обращения: 02.03.2026).

6. Клон и Конъюнктура: МВД РФ анонсировало внедрение ИИ в правоохранительную деятельность к 2025 году // SecurityLab. 2024. URL: <https://www.securitylab.ru/news/545135.php> (дата обращения: 02.03.2026).

7. Об утверждении Стратегии противодействия экстремизму в Российской Федерации : указ Президента Российской Федерации от 28 декабря 2024 г. № 1124 // Гарант.ру. URL: <https://www.garant.ru/products/ipo/prime/doc/411135491/> (дата обращения: 02.03.2026).

8. Харитонов Ю. С. Предвзятость алгоритмов искусственного интеллекта: вопросы этики и права / Ю. С. Харитонов, В. С. Савина, Ф. Паныни // Вестник Пермского университета. Юридические науки. 2021. № 53. С. 488–515.

9. Головин К. С. Необходимость правового регулирования «черного ящика» искусственного интеллекта // Вестник Костромского государственного университета. 2024. Т. 30, № 3. С. 290–297.

10. Ошибка системы Эксмо : когда ИИ не различает слова // РБК. 2026. URL: https://www.rbc.ru/technology_and_media/25/03/2026/69c309d09a7947a5a6fcf550 (дата обращения: 25.03.2026).

11. Об основах государственного регулирования сфер применения технологий искусственного интеллекта в Российской Федерации : проект федерального закона. Доступ из справочно-правовой системы «КонсультантПлюс».

© Рябчиков С. А., 2026

УДК 343.6

Сибгатуллин Ф. Ф. – преподаватель кафедры уголовного права и криминологии Уфимского юридического института МВД России (Россия, г. Уфа)

**БЫТОВОЕ НАСИЛИЕ КАК КРИМИНОЛОГИЧЕСКИЙ
ПРЕДИКТОР ЭКСТРЕМИЗМА
(НА ПРИМЕРЕ СТАТЕЙ 116 И 116.1 УК РФ)**

**DOMESTIC VIOLENCE AS A CRIMINOLOGICAL PREDICTOR
OF EXTREMISM (USING ARTICLES 116 AND 116.1
OF THE CRIMINAL CODE OF THE RUSSIAN FEDERATION)**

Аннотация. В статье исследуется малоизученная в отечественной криминологической науке проблема корреляции между насильственными действиями в семейно-бытовой сфере, ответственность за которые предусмотрена ст.ст. 116 и 116.1 УК РФ, и последующей радикализацией личности, ведущей к формированию экстремистского типа поведения и, в отдельных случаях, к совершению преступлений террористической направленности. Актуальность исследования обусловлена необходимостью расширения традиционного взгляда на детерминацию экстремизма, который преимущественно фокусируется на политических, религиозных и социально-экономических факторах, оставляя без должного внимания микроуровневые процессы социальной агрессии. Автор, опираясь на современные криминологические исследования, новейшие доктринальные источники 2025–2026 гг., а также зарубежный эмпирический опыт (на примере британской программы Prevent), обосновывает тезис о том, что побои являются не просто деликтом против здоровья личности, но и индикатором глубокой социальной агрессии, выступающей «питательной средой» для экстремистской идеологии. В работе анализируются проблемы квалификации побоев, их разграничения со смежными составами преступлений (ст. 115, 117 УК РФ), а также спорные вопросы объекта данных посягательств. Особое внимание уделяется криминологическому механизму