

Дата поступления рукописи в редакцию: 21.06.2026 г.
Дата принятия рукописи в печать: 11.07.2026 г.

DOI: 10.24412/2782-3849-2026-7-85-89

ЦАРЬКОВА Евгения Геннадьевна,

ведущий научный сотрудник ФКУ НИИ ФСИН России, кандидат физико-математических наук,

e-mail: university69@mail.ru

ПРАВОВЫЕ ПРИНЦИПЫ ПРИМЕНЕНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В УГОЛОВНО-ИСПОЛНИТЕЛЬНОЙ СИСТЕМЕ РОССИЙСКОЙ ФЕДЕРАЦИИ: КОНЦЕПТУАЛЬНЫЕ ОСНОВЫ И ПРЕДЕЛЫ РЕГУЛИРОВАНИЯ

Аннотация. Статья посвящена формированию системы правовых принципов применения искусственного интеллекта в уголовно-исполнительной системе Российской Федерации (УИС) с учетом действующих стратегических документов по развитию искусственного интеллекта. Актуальность исследования связана с тем, что внедрение ситуационной видеоаналитики, алгоритмической обработки данных и предиктивных моделей в сфере исполнения наказаний развивается быстрее, чем отраслевое закрепление допустимых пределов их использования, порядка контроля и механизмов защиты прав осужденных и сотрудников. В работе предлагается модель нескольких контуров применения ИИ в УИС (охранно-режимного, предиктивно-профилактического, управленческого и реабилитационного) и вводится категория матрицы допустимости алгоритма, связывающей цель применения, состав данных, степень автономности, форму человеческого контроля и процедуру оспаривания. На этой основе формулируется система правовых принципов, которая позволяет согласовать курс на цифровизацию пенитенциарного ведомства с требованиями законности, пропорциональности, объяснимости, недискриминации и приоритета ресоциализации в сфере исполнения наказаний.

Ключевые слова: искусственный интеллект, уголовно-исполнительная система Российской Федерации, алгоритмическая прозрачность, персональные данные, объяснимый ИИ, алгоритмическая дискриминация.

TSARKOVA Evgeniya Gennadievna,

Senior Researcher of the Research Institute of the Federal Penitentiary Service of Russia, Candidate of Physico-mathematical Sciences

LEGAL PRINCIPLES OF ARTIFICIAL INTELLIGENCE APPLICATION IN THE PENAL ENFORCEMENT SYSTEM OF THE RUSSIAN FEDERATION: CONCEPTUAL FOUNDATIONS AND REGULATORY LIMITS

Annotation. The article is devoted to the formation of a system of legal principles for the application of artificial intelligence in the Penal System of the Russian Federation, taking into account the current strategic documents on the development of artificial intelligence. The relevance of the research stems from the fact that the implementation of situational video analytics, algorithmic data processing, and predictive models in the sphere of sentence enforcement is developing faster than the sectoral entrenchment of the permissible limits of their use, the control procedures, and the mechanisms for protecting the rights of convicts and staff. The paper proposes a model of several application contours for AI in the Penal System (security-regime, predictive-preventive, managerial, and rehabilitative) and introduces the category of the algorithm admissibility matrix, which links the purpose of application, the composition of data, the degree of autonomy, the form of human oversight, and the appeals procedure. On this basis, a system of legal principles is formulated, which makes it possible to harmonize the course towards the digitalization of the penitentiary agency with the requirements of legality, proportionality, explainability, non-discrimination, and the priority of resocialization in the sphere of sentence enforcement.

Key words: artificial intelligence, Penal System of the Russian Federation, algorithmic transparency, personal data, explainable AI, algorithmic discrimination.

Национальная стратегия развития определяет искусственный интеллект (ИИ) искусственного интеллекта до 2030 года как одно из приоритетных направлений

государственной технологической политики. Концепция развития уголовно-исполнительной системы Российской Федерации на период до 2030 года, утвержденная распоряжением Правительства РФ № 1138-р от 29 апреля 2021 года, закрепляет курс на цифровизацию, развитие информационной инфраструктуры и повышение эффективности исполнения наказаний. При этом можно отметить определенные проблемные моменты, связанные с разрывом между стратегическим разрешением цифрового развития и отраслевым правовым режимом применения ИИ в учреждениях ФСИН России. Этот разрыв становится особенно заметен тогда, когда цифровые решения перестают быть просто средствами учета и переходят к функциям интеллектуального анализа. Обычная автоматизация фиксирует и передает данные. Искусственный интеллект интерпретирует данные, выделяет паттерны, классифицирует события, формирует прогноз или рекомендацию. В условиях исправительного учреждения это означает, что технология может начать влиять не только на скорость обработки информации, но и на содержание управленческого либо режимного решения [1, С.242]. С этого момента вопрос о цифровизации становится вопросом о допустимых границах алгоритмического участия в реализации публичной власти.

Для УИС такая постановка вопроса имеет самостоятельное значение. Осужденный находится в отношениях институциональной зависимости от администрации учреждения. Любая ошибка, допущенная при оценке его поведения, рисков или статуса, может привести не только к техническому сбою, но и к ухудшению фактических условий содержания, усилению профилактического контроля, ограничению контактов, изменению маршрута ресоциализации либо к иным последствиям, затрагивающим права и законные интересы человека [2, С.330]. Поэтому здесь недостаточно общей формулы о технологической эффективности, возникает объективная необходимость в выработке специального набора правовых ограничителей.

Содержательно технологии искусственного интеллекта в УИС уже связаны с рядом прикладных сфер, первой из которых является ситуационная видеоаналитика. ГОСТ Р 59385-2021 закрепляет понятие ситуационной видеоаналитики как интеллектуального анализа потокового видео для выявления заранее определенных ситуаций и событий. В практическом измерении речь идет о системах, способных обнаруживать скопление людей, падения, конфликты, атипичные перемещения, оставленные предметы и иные события, значимые для режима и безопасности [3, С.189].

Технологически это развитие уже находится в стадии предметного обсуждения, а не абстрактного прогноза.

Следующая сфера – предиктивная аналитика. Здесь речь идет о моделях, которые могут использоваться для прогнозирования рецидивного поведения, риска суицида, аутоагрессии, конфликтов, побегов, нарушений режима или иных событий, представляющих интерес для администрации учреждения. Такая модель принципиально отличается от простой статистики. Она не только описывает прошлые данные, но и выдает вероятностный вывод о будущем поведении конкретного лица либо группы лиц [4, С. 7]. С правовой точки зрения именно здесь возникает наиболее жесткое напряжение между технологией и правом. Вероятностный прогноз не равен установленному факту. Следовательно, он не может автоматически порождать решение, ухудшающее положение человека.

Следующая и рассматриваемых сфер – управленческая. ИИ потенциально применим для распределения кадровых и материальных ресурсов, анализа нагрузок, обработки обращений, выделения аномалий, выявления организационных рисков и поддержки решений внутри учреждения. На первый взгляд эта область менее чувствительна, поскольку касается внутреннего менеджмента. Однако для пенитенциарного ведомства подобные организационные решения не могут быть нейтральными. Изменение графика обходов, перераспределение внимания личного состава, выделение отдельных зон повышенного риска, изменение доступа к программам сопровождения – все это отражается на фактических условиях отбывания наказания. Следовательно, управленческий ИИ также нуждается в правовой оценке.

Далее, реабилитационная сфера. Концепция развития УИС на период до 2030 года закрепляет ресоциализацию, социальную адаптацию и развитие воспитательной, психологической и социальной работы как стратегические направления развития. В этой логике искусственный интеллект может использоваться для персонализации образовательных траекторий, маршрутизации в программы поддержки, раннего выявления кризисных состояний, распределения консультационных ресурсов и анализа результативности социального сопровождения [5, С. 183].

Из рассмотрения перечисленных выше сфер вытекает тезис о том, что применение искусственного интеллекта в УИС не является единым по юридической природе. Разные контуры затрагивают разные права, создают

специфические риски и требуют разных форм контроля. Ошибка видеоаналитики, пропустившей падение, и ошибка предиктивной модели, ошибочно отнесшей осужденного к группе высокого риска, не совпадают ни по последствиям, ни по юридической оценке. Поэтому единая разрешительная формула для всего ИИ в уголовно-исполнительной системе Российской Федерации невозможна. Существует потребность в дифференцированной модели допустимости.

Такой моделью может выступать матрица допустимости ИИ-алгоритма, используемого в деятельности ФСИН России. Ее смысл состоит в том, что любая система ИИ перед внедрением в УИС должна проходить правовую квалификацию по ряду параметрам. Первый параметр – цель применения. Она должна быть сформулирована точно: обнаружение падений, выявление драки, сортировка обращений, прогноз кризисного поведения, подбор программы сопровождения и т.п. Размытая формула «повышение эффективности» не создает правовой определенности. Следующий параметр – данные. Важно заранее установить, какие категории информации используются: видеопоток, биометрические признаки, сведения о взысканиях, результаты психологического тестирования, медицинские данные, социальные характеристики. Важен также параметр, определяющий степень автономности. Система может только сигнализировать, выдавать рекомендацию либо инициировать действие. Следующий параметр – человеческий контроль. Должно быть ясно, кто именно проверяет вывод алгоритма, в какой срок и под какую ответственность. И, наконец, параметр, определяющий процедуру оспаривания. Если результат ИИ повлиял на правовое положение человека, должно существовать средство проверки и пересмотра такого влияния.

Введение матрицы допустимости обеспечивает последовательное выстраивание правовой логики. Если система не имеет точно определенной цели, она не может быть допущена к использованию. Если цель сформулирована, но состав данных избыточен, вмешательство оказывается неправомерным. Если данные допустимы, но система действует автономно там, где требуется индивидуализированное правовое суждение, возникает запрет на автоматизацию. Если человеческий контроль носит формальный характер, решение становится непрозрачным. Если отсутствует оспаривание, алгоритм превращается в скрытый источник власти. Таким образом, каждый следующий элемент логически вытекает из предыдущего.

Матрица допустимости служит

формированию системы правовых принципов применения ИИ в деятельности сотрудников ФСИН России.

Во-первых, принцип специальной нормативной определенности. Требуется специальный отраслевой нормативный массив, который прямо определит перечень допустимых задач, пределы автономности систем, состав данных и формы контроля. Без этого возникает риск того, что применение ИИ в деятельности ведомства будет развиваться фрагментарно.

Далее, принцип целевой ограниченности. Искусственный интеллект может использоваться только для тех задач, которые прямо закреплены в документации информационной системы и в правовом акте, регулирующем ее внедрение. Так, например, система, предназначенная для обнаружения падений и драк, не должна без нового нормативного основания использоваться для скрытого профилирования личности.

Следующий принцип определяется как сохранение человеческого фактора. Важно, чтобы ИИ и связанные цифровые технологии не заменяли сотрудника, а помогали ему, а решения, затрагивающие права человека, подлежали человеческому пересмотру. Для УИС это означает более жесткое правило, состоящее в том, что окончательные решения, способные ухудшить положение осужденного, не могут приниматься исключительно на основании алгоритма. В ряде случаев должен использоваться коллегиальный порядок проверки, особенно если речь идет о рисках суицида, конфликта, усиленного контроля или иных чувствительных мерах.

Далее, принцип алгоритмической прозрачности и объяснимости. Любая ИИ-система, результаты которой используются в пенитенциарной деятельности, должна оставлять проверяемый цифровой и юридический след. Это означает ведение журнала срабатываний, фиксацию версии модели, регистрационного номера события, ключевых факторов, повлиявших на вывод, уровня достоверности и отметки о проверке сотрудником. Без этих элементов ни прокурорский надзор, ни ведомственный контроль, ни судебное рассмотрение не смогут реконструировать механизм принятия решения.

Следующий принцип – право на информирование и оспаривание. Если алгоритмический вывод использован при принятии или подготовке решения, затрагивающего положение осужденного либо сотрудника, сам факт такого использования должен быть документирован. Не требуется, чтобы человек получал технический доступ к коду системы, но он должен иметь возможность узнать, что его положение оценивалось с использованием ИИ, и добиваться пересмотра

результата через предусмотренную процедуру.

Шестой принцип – минимизация данных. Для каждой задачи должен быть установлен строго ограниченный перечень допустимых данных. Так, для обнаружения драки достаточно видеопотока и сопутствующих технических метаданных. Для оценки кризисного состояния могут потребоваться дополнительные психологические или медицинские сведения, но только при наличии специального правового основания и режима повышенной защиты. Логика «чем больше данных, тем точнее результат» в сфере исполнения наказаний правовой силы не имеет, поскольку избыточный сбор данных прямо повышает риск вторжения в частную сферу.

Далее, принцип недискриминации. Предиктивные и классификационные модели должны проверяться на наличие устойчивых смещений в отношении отдельных групп осужденных. Если алгоритм систематически дает неблагоприятный результат для определенных категорий лиц при сопоставимых фактических обстоятельствах, он не может считаться правомерным инструментом управления. Нормативное требование об антидискриминационной проверке должно предшествовать расширению практики внедрения ИИ-систем.

Вводимый также принцип верифицируемости и аудита обозначает, что система ИИ в УИС должна иметь паспорт, в котором фиксируются разработчик, версия, цели применения, допустимые сценарии использования, показатели точности, ограничения, состав данных и результаты внешней оценки. Только при наличии такого паспорта можно говорить о правомерном включении системы в механизм публичного управления. Иначе контроль будет направлен не на реальную технологию, а на ее декларативное описание.

Не менее важен принцип пропорциональности. Применение ИИ допустимо лишь тогда, когда поставленная цель не может быть достигнута менее инвазивным способом при сопоставимом уровне безопасности либо реабилитационного эффекта. Если задача решается средствами обычной автоматизации, применение биометрии, непрерывного распознавания поведения или глубокой предиктивной аналитики оказывается избыточным. Для пенитенциарной сферы этот тест обязателен, поскольку сама среда исполнения наказания уже характеризуется повышенным объемом ограничений.

И, наконец, принцип распределенной ответственности. Алгоритмическая ошибка не должна растворяться между разработчиком,

поставщиком, оператором и должностным лицом. Разработчик отвечает за архитектуру и обучающие данные, поставщик – за внедрение и сопровождение, учреждение – за соблюдение регламента и журналирование, должностное лицо – за законность итогового решения. Иное решение породило бы правовую фикцию «объективного алгоритма», освобожденного от человеческой ответственности.

Логическим завершением этой конструкции выступают пределы правового регулирования. Первый из этих пределов состоит в том, что личность осужденного не сводится к массиву данных и вероятностному профилю. Алгоритм не может заменить индивидуализированную оценку поведения, мотивации и динамики исправления. Следующий предел выражается в запрете на автоматизацию решений, ухудшающих правовое или фактическое положение лица. Третий предел связан с недопустимостью избыточного сбора данных ради самой технологической возможности их собрать. И, наконец, четвертый предел заключается в том, что интересы безопасности учреждения не устраняют требований законности, соразмерности и уважения человеческого достоинства.

Из этого вытекает практический вывод для нормотворчества. В настоящее время существует объективная потребность в разработке отраслевого акта, регулирующего использование ИИ в деятельности ФСИН России, который должен строиться не как перечень технологий, а как перечень допустимых юридических сценариев. В таком документе целесообразно закрепить определения, закрытый перечень задач по четырем контурам, запрет на полностью автоматические решения, ухудшающие положение осужденного, правила документирования, порядок аудита, режим защиты данных и процедуру оспаривания. Указанный подход согласуется с уже действующими стратегическими документами, но переводит их на более конкретный отраслевой уровень.

Таким образом, логика правового регулирования ИИ в деятельности УИС должна быть развернута не от технологии к ее удобству, а от правового статуса человека к допустимому объему алгоритмического участия в управлении. В первую очередь, важно определить, какое право затрагивается. Затем необходимо выяснить, действительно ли для этой цели нужен ИИ. После этого устанавливаются состав данных, степень автономности, форма человеческого контроля и возможность оспаривания. Только после прохождения этих ступеней система может считаться допустимой [6, С.378]. Именно

такая последовательность позволяет превратить стратегический курс на цифровизацию УИС и развитие ИИ до 2030 года в юридически устойчивую и проверяемую модель правового регулирования.

Список литературы:

[1] Ковалев С.Д. Перспективы использования искусственного интеллекта в исправительных учреждениях ФСИН России // Аграрное и земельное право. 2025, № 1. С. 241-243.

[2] Пекарева В.В., Фроловская Ю.И. Искусственный интеллект в уголовно-исполнительной системе // Право и государство: теория и практика. 2024, № 10. С. 329-331.

[3] Помогалова Ю.В. Зарубежный опыт использования технологии искусственного интеллекта в контексте пенитенциарной безопасности // Вестник Воронежского института ФСИН России. 2024, № 3. С. 188-194.

[4] Андреева Е.А., Беляев А.К., Еремин Ю.М., Царькова Е.Г. Методы биометрической идентификации на основе применения нейросетевых технологий // Информационные технологии в УИС. 2019. № 3. С. 2-10.

[5] Забайкалов А.П. О регламенте Европейского Союза об искусственном интеллекте // Духовно-нравственные ценности как основа технологического суверенитета России: Материалы XVIII Всероссийской научно-практической конференции, Липецк, 24 апреля 2024 г. – Воронеж: Наука-Юнипресс, 2024. С. 181-185.

[6] Царькова Е.Г., Бодров Е.Н. К вопросу

формирования цифровой культуры в процессе профессиональной подготовки сотрудников уголовно-исполнительной системы Российской Федерации // Евразийский юридический журнал. 2023. № 3(178). С. 377-379.

References:

[1] Kovalev S.D. Prospects for the use of artificial intelligence in correctional institutions of the Federal Penitentiary Service of Russia//Agrarian and land law. 2025, № 1. S. 241-243.

[2] Pekareva V.V., Frolovskaya Yu.I. Artificial intelligence in the penal system//Law and the state: theory and practice. 2024, № 10. S. 329-331.

[3] Pomogalova Yu.V. Foreign experience in using artificial intelligence technology in the context of penitentiary security//Bulletin of the Voronezh Institute of the Federal Penitentiary Service of Russia. 2024, № 3. S. 188-194.

[4] Andreeva E.A., Belyaev A.K., Eremin Yu.M., Tsarkova E.G. Methods of biometric identification based on the use of neural network technologies/// Information technologies in UIS. 2019. № 3. S. 2-10.

[5] Zabaykalov A.P. On the regulations of the European Union on artificial intelligence//Spiritual and moral values as the basis of the technological sovereignty of Russia: Materials of the XVIII All-Russian Scientific and Practical Conference, Lipetsk, April 24, 2024 - Voronezh: Nauka-Unipress, 2024. S. 181-185.

[6] Tsarkova E.G., Bodrov E.N. On the formation of digital culture in the process of training employees of the penitentiary system of the Russian Federation//Eurasian Legal Journal. 2023. № 3(178). S. 377-379.



ЮРКОМПАНИ
www.law-books.ru

Юридическое издательство «ЮРКОМПАНИ»
издает научные журналы:

- Научно-правовой журнал «Образование и право», рекомендованный ВАК Министерства науки и высшего образования России (специальности 12.00.01, 12.00.02), выходит 1 раз в месяц.
- Научно-правовой журнал «Право и жизнь», рецензируемый (РИНЦ, E-Library), выходит 1 раз в 3 месяца.