

Дата поступления рукописи в редакцию: 05.05.2026 г.

DOI: 10.24412/2782-3849-2026-6-99-104

Дата принятия рукописи в печать: 25.06.2026 г.

ЦАРЬКОВА Евгения Геннадьевна,

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

e-mail: university69@mail.ru

КОНЦЕПТУАЛЬНЫЕ ОСНОВЫ ВНЕДРЕНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ДЕЯТЕЛЬНОСТЬ УГОЛОВНО-ИСПОЛНИТЕЛЬНОЙ СИСТЕМЫ РОССИЙСКОЙ ФЕДЕРАЦИИ

Аннотация. Применение технологий искусственного интеллекта (ИИ) в уголовно-исполнительной системе Российской Федерации (УИС) формируется преимущественно как технико-организационный тренд без закреплённой нормативной модели и единых правовых ограничений. Это усиливает риск фрагментарного внедрения алгоритмических решений в режимных учреждениях, прежде всего в сферах прогнозирования нарушений установленного порядка отбывания наказания, анализа данных об осуждённых и эксплуатации интеллектуальных систем видеонаблюдения. В статье изложен подход к дифференциации базовых понятий («искусственный интеллект», «робот», «нейронные сети», «большие данные», «искусственная интеллектуальная система»), разработанной в современной доктрине информационного права, и адаптированной к специфике ФСИН России. Формируется концепция «слабого» ИИ как единственно допустимого уровня автоматизации в УИС. Предлагается система правовых принципов внедрения ИИ в деятельность ФСИН России, включая алгоритмическую прозрачность, сохранение «человека в контуре» для решений, затрагивающих права осуждённых, риск-ориентированный дифференцированный режим регулирования для аналитических, рекомендательных и автономных аппаратных систем. Делается вывод о целесообразности разработки ведомственного акта ФСИН России о порядке экспериментального внедрения систем ИИ на базе предложенной системы принципов.

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

CONCEPTUAL FOUNDATIONS FOR THE IMPLEMENTATION OF ARTIFICIAL INTELLIGENCE IN THE ACTIVITIES OF THE PENAL ENFORCEMENT SYSTEM OF THE RUSSIAN FEDERATION

Annotation. The use of artificial intelligence (AI) technologies in the Penal System of the Russian Federation is evolving primarily as a technical and organizational trend, lacking an established regulatory model and uniform legal constraints. This increases the risk of fragmented deployment of algorithmic solutions in correctional institutions, particularly in the areas of predicting violations of the established order of serving sentences, analyzing data on convicted persons, and operating intelligent video surveillance systems. The article presents an approach to differentiating basic concepts («artificial intelligence», «robot», «neural networks», «artificial intelligence system») developed in the modern doctrine of information law, and adapts it to the specific context of the Federal Penitentiary Service of Russia. It formulates the concept of «weak AI» as the only acceptable level of automation within the penal enforcement system. A system of legal principles for the implementation of AI in the Federal Penitentiary Service of Russia is proposed, including algorithmic transparency, maintaining a «human in the loop» for decisions affecting the rights of convicted persons, and a risk-based differentiated regulatory regime for analytical, advisory, and autonomous hardware systems. The article concludes that it is advisable to develop a departmental act of the Federal Penitentiary Service of Russia governing the experimental implementation of AI systems based on the proposed system of principles.

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

За последнее десятилетие эффекты внедрения решений в области искусственного интеллекта (ИИ) вышли далеко за рамки отдельных отраслей и начали радикально менять структуры управления, включая правоохранительную сферу. Сценарий массового использования навигационных сервисов с элементами ИИ, который аккумулирует данные тысяч пользователей для оптимизации маршрутов в режиме реального времени, иллюстрирует базовую идею: алгоритм способен перерабатывать массивы данных, недоступные индивидуальному человеческому наблюдению, и возвращать управленческое решение, воспринимаемое как «оптимальное». Перенос подобной логики в пенитенциарный контекст означает возможность конструирования систем прогнозирования нарушений режима, оптимизации расстановки постов и планирования мероприятий по обеспечению безопасности в учреждениях ФСИН России [1, С.241; 2, С.330].

В зарубежной практике предиктивной аналитики преступности показательным примером является система «Series Finder», разработанная в США для анализа зарегистрированных эпизодов краж и выявления серийных преступлений на основе статистики полиции. Модель, обученная на данных о способе проникновения, типе жилья, времени и месте совершения краж, позволила авторам проекта выделить девять устойчивых паттернов преступного поведения и, как показано в эксперименте, не только восстановить известные серии, но и идентифицировать девять ранее не зафиксированных как взаимосвязанных эпизодов [3, С.190]. Для ФСИН России аналогичная архитектура прогностической системы может использовать данные о нарушениях правил внутреннего распорядка, конфликтах, изъятиях запрещенных предметов, ежедневной планировке маршрутов обхода и особенностях инфраструктуры колонии. В практических терминах это означает построение тепловых карт режимных рисков по локальным зонам учреждения, что позволяет перераспределять посты, усиливать наблюдение в конкретные временные интервалы и обосновывать ресурсные решения по инженерно-техническому укреплению.

Блок потенциальных направлений применения ИИ в деятельности ФСИН России включает несколько типов задач. Первым типом задач является прогнозирование противоправного поведения, где прототипом выступают криминалистические системы типа «Series Finder», но в случае пенитенциарного

ведомства целью такой системы может являться прогнозирование деструктивного поведения осужденных, включая попытки побегов, групповые неповиновения, массовые нарушения режима или актов самоагрессии [4, С.8]. Следующий тип включает задачу оптимизации расстановки сил и средств, решение которой может опираться на алгоритмическую обработку данных о планировке объектов, истории инцидентов и персональном составе отрядов, что позволяет формировать риск-ориентированные графики несения службы, а не статичные штатные расписания. Далее, задача обработка больших данных о личности осужденного, включая сведения о судимостях, дисциплинарных взысканиях, участии в программах ресоциализации, что создает основу для индивидуализации исполнения наказания и оценки рецидивного риска по аналогии с зарубежными инструментами типа COMPAS, являющегося программой для расчета скоринга [5, С.183]. Следующий тип включает видеонаблюдение с элементами ИИ (распознавание лиц, детекция атипичного или агрессивного поведения) и системы поддержки принятия решений при рассмотрении обращений и жалоб осужденных [6, С.377]. Такие задачи образуют самостоятельный класс, где алгоритмы оперируют визуальными и текстовыми потоками данных.

Текущее состояние правового регулирования ИИ в России характеризуется сочетанием точечных технических стандартов и рамочного политико-правового документа в виде Национальной стратегии развития искусственного интеллекта на период до 2030 г., где ИИ определяется как «комплекс технологических решений, позволяющий имитировать когнитивные функции человека...», при этом термин «комплекс технологических решений» фактически описывает «технология искусственного интеллекта», а не сам объект, подлежащий правовой квалификации. ГОСТ Р 59276-2020 и ГОСТ Р 59277-2020 предлагают более точную техническую дефиницию, акцентируя способность системы имитировать когнитивные функции человека, включая самообучение и поиск решений без заранее заданного алгоритма, и связывая ее с комплексом программно-аппаратных средств. Для деятельности ФСИН России такое смещение к программно-аппаратной трактовке ИИ имеет практическое следствие, поскольку объектом правового регулирования и ведомственного учета должна быть не абстрактная «технология», а конкретная система, установленная на

территории учреждения и воздействующая на режим содержания.

Отсутствие специального уголовно-исполнительного регулирования ИИ формирует ситуацию, которую можно обозначить как «правовой коллапс», представляющий собой состояние, при котором понятие искусственного интеллекта выступает понятием, в которое различные акторы вкладывают несопоставимое содержание. При этом ожидание окончательного прояснения технической природы ИИ ведут к нарастанию непрозрачных рисков, а не к их снижению. Для ФСИН России такая ситуация означает, что попытки использовать ИИ-решения в пилотном режиме (например, при установке интеллектуальных камер или внедрении систем риск-скоринга осужденных) будут происходить в условиях отсутствия базовых понятий и критериев приемлемости ошибок. Соответственно, минимизация правовых рисков требует не отложения регулирования, а выстраивания системы принципов, способной охватить различающиеся по задачам классы систем без навязывания единого универсального определения.

Ключевым элементом методологического блока данного исследования является разграничение ИИ и смежных технико-правовых понятий. Искусственный интеллект может быть определен как комплекс программно-аппаратных средств, обеспечивающий полноту или фрагмент информационного цикла: восприятие, обработку, хранение информации и формирование решения. Необходимым признаком технологии выступает либо зависимость конечного решения от человека (слабый ИИ), либо автономность (сильный ИИ), причем сильный ИИ в существующей технологической парадигме пока не реализован.

Зачастую робот интерпретируется как воплощенный искусственный интеллект, однако он может функционировать и на основе элементарных алгоритмов без использования технологий ИИ, что имеет прямое значение для деятельности учреждений ФСИН России при внедрении робототехнических систем (роботы-патрульные, беспилотные летательные аппараты) поскольку формальное наличие «робота» не гарантирует применения ИИ, а следовательно, не требует специальных процедур валидации и аудита моделей.

Нейронные сети и методы машинного обучения представляют собой интеллектуальные средства, обеспечивающие итерационный поиск оптимума некоторого функционала качества решения задачи. Нейронная сеть реализуется в форме многомерной нелинейной системы, которая в ходе обучения подстраивает свои параметры с целью минимизации ошибки

прогноза, а глубокое обучение и машинное обучение служат дисциплинами, изучающими построение самообучающихся алгоритмов на основе разных режимов (с учителем, без учителя, с подкреплением). В контексте внедрения искусственного интеллекта в деятельность учреждений ФСИН России внедрение данных технологий означает, что каждое применение нейросетевой модели, например, к оценке риска побега или агрессивного поведения привязано к конкретной выборке исторических данных учреждения, а следовательно, зависит от полноты и релевантности этой выборки. При этом комбинация нейронных сетей и машинного обучения не ведет к появлению сильного ИИ, поскольку подобные архитектуры решают конечные, формализуемые задачи, что снимает часть возражений о замене персонала исправительных учреждений алгоритмами.

Тезис о наличии правового коллапса подводит к выводу об отказе от поисков единого определения ИИ в пользу отраслевых дефиниций, учитывающих специфику соответствующих правоотношений. Правовая регламентация ИИ не может откладываться до момента окончательного согласования технических терминов, поскольку речь идет о рынке, который еще не сформировался, но уже создает экзистенциальные риски. Применительно к деятельности УИС это означает, что основные параметры допустимого использования ИИ, включая виды решаемых с помощью ИИ задач, допустимые уровни ошибок, требование участия человека в принятии решений, должны быть закреплены в уголовно-исполнительном законодательстве и ведомственных актах еще до массового внедрения конкретных систем. Универсальное широкое определение ИИ целесообразно сохранять на уровне федерального закона об искусственном интеллекте и Национальной стратегии, тогда как для ФСИН России целесообразна функциональная дефиниция, ориентирующаяся на слабый ИИ и конкретные прикладные задачи пенитенциарного ведомства.

С учетом вышеизложенного, целесообразно закрепление на уровне федерального закона общего понятия ИИ как совокупности информационных технологий, позволяющих решать задачи, сопоставимые с человеческим интеллектом, а технические характеристики и классификации оставлять на уровне ГОСТ и подзаконных актов. Преимущество такой модели для пенитенциарной сферы состоит в гибкости, поскольку в таком случае ведомство может обновлять перечень видов ИИ-систем, процедуры их тестирования и показатели качества через приказы и

инструкции без постоянного внесения изменений в уголовно-исполнительный кодекс. Для строгой научной формулировки достаточно указать, что широкое определение ИИ в сочетании с техническим регулированием создает многоуровневую модель, где федеральный закон задает рамки, а ведомственные акты – конкретные режимы. Может быть предложена следующая совокупность уточнений определения ИИ при его использования для решения пенитенциарных задач. Во-первых, термин «комплекс технологических решений» целесообразно заменить на «комплекс программно-аппаратных средств». Далее, в определение следует включить указание на «совокупность интеллектуально-волевых средств и способов взаимодействия с окружающей действительностью», подчеркивая наличие не только инфраструктуры обработки данных, но и автономных механизмов выбора вариантов решения внутри заданных человеком целей. Это позволит прямо зафиксировать, что объектом регулирования ИИ, применяемого в интересах ФСИН России являются установленные в пенитенциарных учреждениях программно-аппаратные комплексы, способные имитировать когнитивные функции человека и формировать выводы, используемые персоналом при оценке поведения осужденных и принятии управленческих решений.

Сопоставление российских подходов с европейским регулированием демонстрирует дополнительный аргумент в пользу риск-ориентированного, а не функционально-нейтрального подхода к ИИ. В документе «Artificial Intelligence for Europe» (2018 г.) ИИ определяется как система, демонстрирующая разумное поведение при анализе окружающей среды и действиях с определенной степенью автономности для достижения конкретных целей, при этом прямо оговаривается возможность как программной, так и аппаратной реализации. Принятый в 2024 г. Закон ЕС об искусственном интеллекте (EU AI Act) вводит категорию систем высокого риска, к которым относятся, в частности, ИИ-системы, используемые в правоохранительной и пенитенциарной деятельности, и предъявляет к ним повышенные требования по управлению данными, документации и человеческому надзору. Для пенитенциарных учреждений это означает, что любая попытка использования ИИ в оценке риска побега, классификации осужденных по режимам или автоматизированном принятии решений о переводе в иные условия отбывания наказания должна исходно рассматриваться как работа с системой высокого риска, а не как нейтральная цифровая автоматизация.

В ходе исследования выработана система правовых принципов и ограничений внедрения ИИ в УИС. Ключевой принцип состоит в алгоритмической прозрачности пенитенциарных ИИ-систем. Это означает, что любая система ИИ, используемая в учреждениях ФСИН России, должна иметь документированную архитектуру, описание входных и выходных данных, методологию обучения и показатели качества, которые подлежат ведомственному контролю и могут быть проверены при рассмотрении жалоб и обращений осужденных.

Далее, принцип сохранения «человека в контуре» (human-in-the-loop) для всех решений, затрагивающих права осужденных. С учетом невозможности реализации эмоционального интеллекта ИИ может использоваться только как инструмент поддержки, но не замены усмотрения должностного лица. Поэтому требует прямого закрепления в УИК РФ и ведомственных актах, тезисов означающих, что решения о переводе в облегченные или строгие условия, применении взысканий, поощрений, подготовке материалов об условно-досрочном освобождении принимаются комиссионно или индивидуально человеком, который несет юридическую ответственность и не вправе ссылаться на ИИ как на единственное основание. Функция ИИ сводится исключительно к предоставлению аналитических материалов, а итоговая юридически значимая оценка принадлежит сотруднику, прошедшему специальную подготовку по интерпретации выводов моделей.

Следующий принцип включает дифференцированный правовой режим для трех типов ИИ-систем во ФСИН России. Предлагается выделить следующие типы прикладных инструментов: аналитические системы (для построения прогноза нарушений, криминогенной картографии учреждений), системы поддержки принятия решений (для скоринга рецидивного риска, классификации и обработки обращений, анализа видеопотока для выдачи тревожного сообщения на пост охраны), автономные аппаратные комплексы (роботизированные платформы охраны периметра, автономные средства доставки внутри режимных зон и др.). Для первой группы достаточно предусмотреть требования к качеству данных, валидации моделей на исторических выборках и регулярному пересмотру параметров, поскольку прямое воздействие на права осужденных здесь опосредовано через последующие решения персонала. Для второй группы требуется усиление режима, включая обязательное участие человека, запрет на полную автоматизацию ключевых решений, фиксирование в делах отметок о том, была ли использована ИИ-система и как

оценена ее рекомендация. Третья группа должна рассматриваться как источник повышенной опасности по аналогии с транспортными средствами и промышленными объектами, что влечет особый уровень ответственности за ущерб, причиненный осужденным, личному составу и третьим лицам.

Четвертый принцип можно условно обозначить как принцип функциональной дифференциации дефиниций. Вместо попытки закрепить в УИК РФ универсальное понятие ИИ предлагается определить на уровне уголовно-исполнительного законодательства несколько юридически значимых категорий: «система искусственного интеллекта для анализа пенитенциарных данных», «система искусственного интеллекта для технической защиты и охраны объектов ФСИН России», «система искусственного интеллекта, используемая в воспитательной и психологической работе». Для каждой категории формулируются отдельные ограничения и требования, включая перечень допустимых задач, предельные уровни доверия к рекомендациям и обязательные процедуры аудита. Такой подход позволяет избежать ситуации, когда единая дефиниция становится либо чрезмерно широкой, либо неприменимой к конкретным пенитенциарным практикам.

Отдельного рассмотрения требует вопрос о соотношении информационного обеспечения и ИИ в деятельности ФСИН России. Информационное обеспечение правоприменения понимается как совокупность данных, процедур их сбора, обработки, хранения и предоставления пользователю, обеспечивающая возможность принятия обоснованных решений. ИИ, в этой логике, выступает не самостоятельным источником информации, а инструментом, который обеспечивает более сложную, многомерную обработку уже имеющихся данных и формирование новых производных показателей. Для УИС это означает, что внедрение ИИ должно опираться на предварительное упорядочение базовых массивов, от электронных личных дел осужденных до реестров режимных инцидентов, и не может компенсировать пробелы в первичной регистрации. Иными словами, алгоритм накладывается на систему информационного обеспечения, но не заменяет ее.

С точки зрения организационного оформления набор принципов способен лечь в основу ведомственного акта, условно обозначаемого как приказ «О порядке экспериментального внедрения систем искусственного интеллекта в учреждениях УИС». В такой акт могут быть включены положения, связанные с классификацией ИИ-систем по

трем типам, предложенным выше, требования к данным и процедурам валидации, порядок включения человека в контур принятия решений и документирования отношения должностного лица к рекомендациям ИИ, тезисы по регламентации аудита алгоритмов и внешней экспертизы систем высокого риска, критерии прекращения эксперимента при достижении определенных порогов ошибок или выявлении дискриминационных эффектов. При разработке такого акта может быть использована риск-ориентированная структура EU AI Act с адаптацией к отечественным реалиям и уголовно-исполнительной специфике.

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.

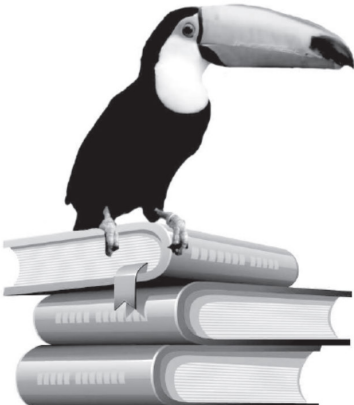
[3] Pomogalova Yu.V. Foreign experience in using artificial intelligence technology in the context of penitentiary security//Bulletin of the Voronezh Institute of the Federal Penitentiary Service of Russia. 2024, № 3. S. 188-194.

[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.



**ЮРИДИЧЕСКОЕ ИЗДАТЕЛЬСТВО
«ЮРКОМПАНИ»**

Издание учебников,
учебных и методических
пособий, монографий,
научных статей.

Профессионально.

В максимально
короткие сроки.

Размещаем
в РИНЦ, E-Library.

ЮРКОМПАНИ

www.law-books.ru