Этико-правовые риски внедрения технологий искусственного интеллекта в процессе цифровизации государственного управления

Панина Елена Николаевна
кандидат психологических наук, профессор кафедры государственно-правовых дисциплин, Сибирский юридический институт МВД России, conf_22@inbox.ru

В статье исследуются этико-правовые риски внедрения технологий искусственного интеллекта в процессе цифровизации государственного управления. Цель работы состоит в выделении и систематизации рисков, возникающих при применении интеллектуальных систем в деятельности органов публичной власти. Обосновывается, что использование искусственного интеллекта в публичной сфере отличается от его применения в коммерческом секторе, поскольку алгоритмически поддерживаемые решения способны непосредственно влиять на реализацию конституционных прав, публично-правовой статус граждан и доверие к государственным институтам. Выделены риски алгоритмической непрозрачности, дискриминации, ошибочной обработки данных, нарушения приватности, размывания публичной ответственности, цифрового неравенства, технологической зависимости от разработчиков, киберугроз, неправомерного профилирования и подмены индивидуального рассмотрения автоматизированной оценкой. Каждый риск рассмотрен через возможные сферы применения ИИ в государственном управлении: предоставление государственных услуг, социальная поддержка, контрольно-надзорная деятельность, здравоохранение, миграционный контроль, государственные закупки и правоохранительная деятельность. Делается вывод о необходимости разработки единой модели искусственного интеллекта в публичной сфере, включающей принципы и механизмы человекоцентричного использования технологий: законность, приоритет прав и свобод, человеческий контроль, объяснимость решений, подотчетность, недискриминацию, минимизацию данных, аудит алгоритмов и доступность внесудебного и судебного обжалования.

Ключевые слова: искусственный интеллект; государственное управление; цифровизация; этика искусственного интеллекта; права человека; алгоритмическая прозрачность; цифровое неравенство; публичная ответственность; человекоцентричный подход; государственные услуги.

Введение

Внедрение технологий искусственного интеллекта в государственное управление является закономерным направлением цифровой трансформации публичного сектора. Использование интеллектуальных систем позволяет автоматизировать документооборот, обрабатывать обращения граждан, осуществлять межведомственный обмен сведениями, прогнозировать риски, выявлять аномалии в государственных закупках, анализировать массивы статистических данных и поддерживать принятие управленческих решений. Национальная стратегия развития искусственного интеллекта на период до 2030 года связывает развитие ИИ с решением задач социально-экономического развития, совершенствованием государственного управления, повышением доступности услуг и укреплением технологического потенциала Российской Федерации (Указ Президента Российской Федерации от 10 октября 2019 г. № 490 «О развитии искусственного интеллекта в Российской Федерации»).

Вместе с тем применение ИИ в публичной сфере имеет принципиально иное значение, чем его использование в коммерческой деятельности. В частном секторе алгоритмическая ошибка обычно связывается с экономическими потерями, репутационными издержками либо снижением качества сервиса. В государственном управлении последствия могут выражаться в необоснованном отказе в государственной услуге, ограничении доступа к социальной поддержке, проведении проверки, нарушении права на неприкосновенность частной жизни, ошибочном отнесении лица к группе риска или фактическом ограничении возможности реализовать субъективное право.

ИИ в публичном управлении действует не изолированно от правовой системы. Он включается в процедуры реализации компетенции государственных органов, подготовки административных решений, распределения публичных ресурсов и осуществления контроля. Поэтому применение таких технологий должно оцениваться не только через критерии технологической точности и экономической эффективности, но и с точки зрения принципов правового государства, законности, равенства, справедливости, юридической определенности, защиты достоинства личности и гарантированности судебной защиты.

Научная новизна настоящего исследования заключается в выделении и систематизации этико-правовых рисков, возникающих именно при внедрении технологий искусственного интеллекта в систему государственного управления. В отличие от общего описания угроз цифровизации, рассматриваемые риски соотносятся с конкретными видами публичной деятельности и юридическими последствиями, которые могут наступить для граждан, организаций и органов власти. Такой подход позволяет показать, что проблема ИИ состоит не только в технической надежности алгоритма, но и в сохранении конституционно-правовой природы публичного решения.

Специфика ИИ в публичной сфере

Искусственный интеллект может использоваться в государственном управлении на разных уровнях. С точки зрения наименьшего риска может быть рассмотрено использование ИИ для реализации технических и вспомогательных функций: распознавания тех или иных документов, поиска необходимой информации, составления проектов ответов на запросы граждан, сортировки обращений к должностным лицам, выявления дублирующих сведений и контроля сроков исполнения. В указанных ситуациях технология не определяет как таковое содержание публичного решения, а лишь способствует оптимизации внутренней (технической) деятельности конкретного органа власти.

Применение ИИ в сфере аналитики определяет более высокий уровень этико-правовых рисков. В частности, алгоритм вполне способен выявлять коррупционные риски, формировать конкретные прогнозы, оценивать вероятные нарушения при организации исполнения того или иного управленческого решения, определять приоритетность рассмотрения обращений граждан или направлять ресурсы на определенные территории. Даже если подобная система формально не принимает окончательное решение, ее вывод способен существенно повлиять на принятие решения уполномоченного должностного лица.

И, наконец, максимальный риск присутствует в тех ситуациях, когда ИИ непосредственно участвует в управленческих процедурах,

затрагивающих правовой статус личности. Например, определении права на социальную выплату, принятии окончательного решения о предоставлении либо отказе в государственной услуге, выборе объекта проверки или контроля, оценке миграционного риска, обработке сведений о здоровье субъекта, назначении мер социальной поддержки или выявлении признаков правонарушения. В подобных случаях алгоритмическая рекомендация вполне может приобрести обязательный характер, а должностное лицо превращается в формального оператора-исполнителя автоматизированного решения.

Таким образом, именно публично-правовая природа решений в процессе государственного управления предопределяет особые, повышенные требования к использованию ИИ. Государство в лице должностных лиц не вправе ссылаться на точную определенность алгоритма как на самостоятельное основание для вмешательства в сферу прав граждан. Критерий технологической эффективности не должен заменять критерий законности, а автоматизация, как аспект принятия управленческого решения, не должна исключать необходимость принятия должностным лицом мотивированного решения, гарантирующего равенство участников процесса и возможность его обжалования.

Далее последовательно рассмотрим этико-правовые риски внедрения технологий искусственного интеллекта в сферу государственного управления.

Алгоритмическая непрозрачность

В качестве одного из значимых рисков, по нашему мнению, выступает алгоритмическая непрозрачность. В литературе подобное явление нередко описывается как «черный ящик». Суть его состоит в том, что логику результата, полученного с использованием алгоритма ИИ, сложно, а порой и невозможно, объяснить не только обычному пользователю или должностному лицу, но и разработчику алгоритма.

Очевидно, что в государственном управлении непрозрачность несовместима с требованием мотивированности административного решения. Гражданин должен понимать, по каким причинам ему отказано в выплате, почему его заявление направлено на дополнительную проверку, по каким критериям организация отнесена к категории риска и т.д. Если единственным объяснением становится ссылка на вывод системы, возникает ситуация, при которой право на получение информации о причинах решения становится формальным.

Например, алгоритм в сфере социальной поддержки может не рекомендовать назначение выплаты из-за несоответствия сведений, полученных из разных реестров. Однако данные могут оказаться устаревшими, содержать техническую ошибку или не учитывать конкретные обстоятельства семьи. Если гражданину не разъяснены основания автоматизированного вывода и не предоставлена возможность представить уточняющие документы, цифровой сервис начинает ограничивать, а не обеспечивать доступ к государственной поддержке.

Рекомендация ЮНЕСКО об этических аспектах искусственного интеллекта исходит из необходимости прозрачности и объяснимости систем ИИ, особенно когда их применение затрагивает права и свободы человека (UNESCO. Recommendation on the Ethics of Artificial Intelligence, 2021). В публичной сфере объяснимость означает не обязательное раскрытие всего исходного кода, а предоставление понятной информации о том, какие юридически значимые данные были использованы, какие критерии сыграли определяющую роль и какой субъект несет ответственность за итоговое решение.

Алгоритмическая дискриминация

Следующий этико-правовой риск связан с алгоритмической предвзятостью и, как следствие, дискриминацией. Система ИИ обучается на данных, которые не исключают ошибки, пробелы, а также могут содержать социальные стереотипы и результаты ранее сложившейся практики. Необходимо понимать, что алгоритм не является нейтральным только потому, что он технический, он воспроизводит структуру и качество информации, на которой обучался, формировал критерии принятия решения.

В государственном управлении предвзятость особенно опасна при применении ИИ в таких сферах как: социальная политика, контрольная деятельность, здравоохранение, миграционное администрирование и правоохранительная деятельность. Так, алгоритм, формирующий приоритет проверок, может опираться на сведения о ранее проведенных контрольных мероприятиях и их результатах. Если определенные категории организаций или территории исторически проверялись чаще, система может закрепить этот паттерн и постоянно относить их к группе повышенного риска без достаточных на то объективных оснований.

Аналогичные проблемы возможны при оценке нуждаемости граждан в той или иной услуге. Если алгоритм использует косвенные признаки – место проживания, тип занятости, состав семьи, данные о расходах, – он может давать заниженную оценку права на поддержку лицам, чьи жизненные обстоятельства не укладываются в типовую модель. Формально нейтральные критерии способны производить дискриминационный эффект, если непропорционально затрагивают определенные группы населения.

Конституционный принцип равенства требует, чтобы применение ИИ не приводило к необоснованному различию в обращении с гражданами. Снижению данного риска должна способствовать проверка обучающих данных до внедрения системы, тестирование результатов для разных групп и постоянный мониторинг ошибочных решений. При проявлении устойчивого дискриминационного эффекта алгоритм должен быть скорректирован, а его использование – приостановлено до устранения выявленных нарушений.

Ошибки данных и автоматизированная несправедливость

Третий этико-правовой риск состоит в использовании неполных, недостоверных или устаревших данных. ИИ не способен компенсировать недостатки исходной информации, а, напротив, может масштабировать ошибку, распространяя ее на большое число решений.

Специфика решений в сфере государственного управления состоит в том, что многие из них базируются на сведениях из соответствующих реестров. Например, право на пособие может зависеть от данных о доходах, составе семьи, регистрации по месту жительства, наличии или отсутствии имущества. Если хотя бы один из источников содержит неточную информацию, алгоритм вполне может допустить необоснованный вывод, ошибка может быть воспроизведена без анализа конкретной информации. В случае же принятия решения конкретным должностным лицом, оно может зафиксировать неточность или противоречие и инициировать уточнение информации или запросить дополнительные документы.

Представим ситуацию, когда гражданин меняет место работы, переходит с высокооплачиваемой на деятельность со средней или даже низкой заработной платой. При этом не исключается, что сведения в реестре доходов обновляются с задержкой. Алгоритм, оценивающий право на социальную выплату, использует прежние сведения и автоматически формирует отказ. Таким образом, гражданин сталкивается не только с фактической ошибкой, но и с трудностью ее исправления, если порядок пересмотра итогового решения не предусмотрен или неочевиден.

В данном случае этико-правовой риск заключается в подмене принципа индивидуального рассмотрения дела в конкретной, изменяющейся ситуации формальной обработкой абстрактного цифрового профиля. Очевидно, что для максимального учета прав и интересов граждан государственное управление не может исходить из презумпции безошибочности внесенных в реестры данных. Напротив, оно должно предоставлять гражданину возможность ознакомиться со сведениями, на основании которых принято решение, при необходимости исправить их и инициировать повторное рассмотрение обращения.

Нарушение приватности и чрезмерное наблюдение

Четвертым этико-правовым риском является угроза неприкосновенности частной жизни. Функционирование ИИ предполагает обработку значительных массивов информации, включая персональные данные, сведения о здоровье, финансовом положении, месте нахождения, семейных связях, поведении граждан и т.д. Чем больше данных получает алгоритм, тем точнее может быть его прогноз. Однако увеличение объема данных не должно превращаться при осуществлении государственного управления в самоцель.

Федеральный закон «О персональных данных» закрепляет принципы законности, справедливости, ограничения обработки конкретными законными целями и соответствия содержания и объема данных заявленным целям обработки (Федеральный закон от 27 июля 2006 г. № 152-ФЗ «О персональных данных»). Эти требования в полной мере должны применяться и к системам ИИ. Техническая

возможность собрать и сопоставить данные не создает самостоятельного правового основания для активного вмешательства в частную жизнь граждан.

Например, использование ИИ для выявления лиц, имеющих право на социальную поддержку, может требовать анализа сведений о доходах и составе семьи. Однако сбор данных о перемещениях, интернет-активности или контактах гражданина для этой цели будет несоразмерным. Отсутствие ясных ограничений способно привести к формированию цифрового профиля личности, который используется шире первоначально определенной цели.

Особую угрозу представляет объединение разрозненных баз данных без четко определенной правовой цели и надлежащего контроля доступа. Такая практика увеличивает риск утечки, неправомерного использования сведений, дискриминационного профилирования и несанкционированного наблюдения. В связи с этим применение ИИ должно быть основано на принципе минимизации запрашиваемых данных, т.е. использование только тех сведений, без которых конкретная публичная задача не может быть решена.

Размывание ответственности

Пятым этико-правовым риском выступает размывание публичной ответственности за принятое решение. При использовании ИИ в процессе принятия решения возможна ситуация, когда должностное лицо ссылается на вывод системы, государственный орган – на технического оператора, а разработчик – на некорректное использование продукта. В результате гражданин не понимает, кто несет ответственность за ошибку и к кому обращаться за ее исправлением.

Такой подход противоречит основам публичной власти: ИИ не является субъектом права, не обладает самостоятельной компетенцией и не может нести юридическую ответственность. Он не может быть адресатом жалобы или ответчиком по административному делу. Следовательно, применение ИИ не должно разрывать связь между принятием публичного решения и ответственным за данное направление государственным органом.

Например, если система ошибочно определила гражданина как не имеющего права на льготу, ответственность за отказ не может быть перенесена на алгоритм. Государственный орган обязан обеспечить качество данных, правомерность критериев, проверку результата и возможность пересмотра. Частный разработчик может нести договорную, гражданско-правовую или иную ответственность в пределах отношений с заказчиком, однако перед гражданином публичную ответственность должен сохранять орган власти.

Для устранения данного риска представляется целесообразным заранее распределять роли, обязанности и ответственность всех участников: заказчика, разработчика, оператора, администратора системы, должностного лица, принимающего итоговое решение, и органа, рассматривающего жалобу. Такое распределение должно быть закреплено в нормативных правовых актах, регламентах и технических заданиях.

Деформация доверия и легитимности

Шестой этико-правовой риск состоит в снижении доверия граждан к публичной власти. Доверие связано не только с результативностью государственного управления, но и с понятностью, справедливостью и доступностью процедур. Если гражданин получает отказ от системы, не может выяснить его причины и не имеет понятного способа оспорить решение, им цифровизация воспринимается как отчуждение государства от человека.

Данная проблема особенно остра в ситуациях использования ИИ для оценки рисков, выявления потенциальных нарушений или ранжирования граждан по различным параметрам. Переход к «управлению по вероятности» может изменить саму природу отношений государства и личности: неблагоприятные последствия наступают не за уже совершенное противоправное деяние, а за прогнозируемую модель поведения.

Например, алгоритм может рекомендовать проведение дополнительной проверки в отношении организации на основе совокупности статистических признаков. Безусловно, статистическая аналитика значительно облегчает принятие управленческого решения, но она должна носить рекомендательный характер. Такая рекомендация допустима как вспомогательная информация, но не должна являться юридическим основанием для проверки. В том случае, если прогноз ИИ становится фактической причиной дополнительной проверки без надлежащего анализа, принцип законности заменяется вероятностной оценкой, что недопустимо при осуществлении государственной управленческой деятельности.

Этот риск особенно опасен тем, что снижение доверия в долгосрочной перспективе может нивелировать преимущества цифровизации в социальном аспекте. Государство, использующее непрозрачные алгоритмы, рискует сформировать внешне современную, но социально нелегитимную систему управления. В связи с этим оптимально открытая информация о применении ИИ, наличие механизмов обратной связи и гарантии обжалования принятого решения являются не дополнительными, а обязательными элементами цифрового публичного управления.

Цифровое неравенство и исключение

Седьмой этико-правовой риск проявляется в цифровом неравенстве. ИИ часто внедряется вместе с цифровыми платформами, электронными идентификаторами, дистанционными формами подачи заявлений и автоматизированными сервисами. Безусловно, для значительной части граждан этот формат повышает удобство и значительно сокращает продолжительность получения услуг. Однако в настоящее время сохраняется значительное количество граждан, для которых цифровизация может стать новым препятствием в процессе получения государственных услуг: лица с недостаточным доступом к интернету, низкой цифровой грамотностью, инвалидностью, языковыми барьерами или проживанием в удаленных территориях.

Если получение государственной услуги базируется только на цифровой платформе, то гражданин, не обладающий соответствующими навыками или техническими средствами, оказывается в невыгодном положении. Это обстоятельство само по себе уже противоречит принципу равного доступа к публичным услугам. Ослабить проявление этого риска должно сохранение, наряду с использованием ИИ, традиционных способов обращения, особенно в сферах социальной защиты, здравоохранения, правосудия и миграционных процедур.

Например, чат-бот, предназначенный для информирования граждан о мерах социальной поддержки, весьма полезен для оперативного ответа на типовые вопросы. Однако он не должен быть единственным каналом коммуникации для пожилого человека, лица с нарушением зрения или гражданина, нуждающегося в индивидуальном разъяснении сложной правовой ситуации. В таких случаях для решения должна сохраняться возможность прямого обращения к человеку – специалисту органа власти или многофункционального центра.

Технологическая зависимость и приватизация публичной функции

Восьмым риском является зависимость государственных органов от частных разработчиков и поставщиков технологий. Большинство сложных ИИ-решений создается коммерческими организациями, которые могут обладать исключительными правами на программное обеспечение, контролировать обновления, архитектуру и отдельные компоненты используемой системы. Если государственный орган не располагает достаточной технической экспертизой, он может оказаться зависимым от поставщика в вопросах, затрагивающих реализацию публичных полномочий.

Такая зависимость создает этический и правовой риск скрытой приватизации публичной функции. Разработчик может влиять на параметры системы, критерии классификации, порядок обработки данных и возможности аудита. При этом гражданин взаимодействует с публичной властью, но фактическая логика обработки его обращения определяется частной инфраструктурой.

Для предупреждения данного этико-правового риска необходимы четко сформулированные требования к государственным контрактам на разработку и внедрение ИИ. Они должны предусматривать право заказчика на аудит, доступ к документации, контроль за обновлениями, обеспечение информационной безопасности, возможность независимой экспертизы и порядок передачи данных. В чувствительных сферах государство должно сохранять возможность самостоятельного контроля над ключевыми компонентами системы.

Киберугрозы и манипулирование результатами

Девятый этико-правовой риск связан обеспечением безопасности ИИ-систем. Алгоритмы могут быть объектом внешних атак, неправомерного изменения данных, подмены обучающей выборки и манипулирования результатами. В публичной сфере подобные атаки способны повлиять на распределение бюджетных средств,

обработку обращений, государственные закупки, реестровые записи и контрольные мероприятия.

Например, если в системе оценки рисков контролируемых субъектов будет изменен набор данных или параметров модели, отдельная организация может быть неправомерно исключена из группы контроля либо, напротив, необоснованно подвергнута дополнительным проверкам. В сфере государственных услуг вмешательство в алгоритм может привести к задержкам выплат, ошибочному отказу или утечке персональной информации.

Федеральный закон «Об информации, информационных технологиях и о защите информации» устанавливает правовые основы защиты информации, а законодательство о критической информационной инфраструктуре предусматривает специальные требования безопасности для значимых объектов (Федеральный закон от 27 июля 2006 г. № 149-ФЗ «Об информации, информационных технологиях и о защите информации»; Федеральный закон от 26 июля 2017 г. № 187-ФЗ «О безопасности критической информационной инфраструктуры Российской Федерации»). Однако применение ИИ требует дополнительного внимания к специфическим угрозам: отравлению данных, атаке на модель, подмене результатов и использованию уязвимостей при интеграции систем.

Неправомерное профилирование

Десятый этико-правовой риск заключается в возможном неправомерном профилировании граждан: ИИ способен объединять разнородные сведения и на этом основании прогнозировать возможные модели поведения как отдельных граждан, так и целых социальных групп. В зависимости от целей подобный анализ может использоваться, с одной стороны, для выявления мошеннических схем, а с другой – для определения права на социальную поддержку или оценку вероятности нарушения правовой нормы или государственного установления. При этом граница между правомерным анализом данных и чрезмерным профилированием является весьма размытой.

Чрезмерное профилирование опасно тем, что создает предпосылки для принятия решений не на основе конкретных юридических фактов, а с опорой на вероятностные характеристики. Гражданин может быть отнесен к категории риска не по факту совершения им нарушения, а в связи с конкретным местом проживания, проведением тех или иных финансовых операций или наличием определенных социальных связей. Подобный подход противоречит принципу индивидуализации публичного решения и способен привести к стигматизации отдельных групп.

Таким образом, при осуществлении цифровизации государственного управления, необходимо учитывать, что применение прогнозных моделей допустимо лишь как вспомогательный инструмент. Оно не может быть единственным основанием для ограничения прав, отказа в услуге, привлечения к ответственности или проведения контрольного мероприятия. В каждом конкретном случае при принятии управленческого решения необходимо устанавливать конкретные юридические факты, а не ограничиваться статистической оценкой вероятности проявления того или иного варианта поведения.

Необходимость единой модели человекоцентричного ИИ

Анализ выделенных рисков позволяет сделать следующий вывод: регулирование использования ИИ в государственном управлении не может сводиться к требованиям технической безопасности или защите персональных данных. Целесообразно внедрение единой модели использования ИИ в публичной сфере, соединяющей этические ориентиры с юридически обязательными процедурами.

При этом необходимо исходить из того, что ИИ является лишь вспомогательным инструментом публичной власти, а не самостоятельным субъектом управления. Для оптимального функционирования подобной модели ее содержание должно определяться следующими принципами:

наличие четко определенной публичной цели, в обязательном порядке предусматривающей законность применения ИИ;

приоритет прав, свобод и достоинства человека при осуществлении публичной управленческой деятельности;

сохранение контроля человека-должностного лица в качестве обязательного аспекта при принятии юридически значимого управленческого решения;

понятность для граждан и возможность оспорить ими алгоритмически поддерживаемые решения;

персональная ответственность государственного органа и соответствующего должностного лица за конкретное принятое управленческое решение;

недопустимость дискриминации и регулярная проверка алгоритмической предвзятости при обеспечении управленческого решения;

минимизация запрашиваемых данных, ограничение целей их обработки и защита приватности в соответствии с конкретным запросом на государственную услугу;

безопасность информационных систем и устойчивость к внешнему вмешательству;

доступность для граждан альтернативных, нецифровых форм взаимодействия с государством;

независимый аудит систем ИИ и общественная подотчетность.

Механизмы реализации данной модели должны включать предварительную оценку воздействия ИИ-системы на права граждан; классификацию систем по уровню риска; обязательный аудит высокорисковых алгоритмов; ведение реестра ИИ-систем, используемых государственными органами; регламентацию процедур контроля качества данных; фиксацию действий пользователей; право гражданина на уведомление об использовании ИИ и на пересмотр решения человеком; создание ведомственных этических комиссий; а также развитие цифровой грамотности государственных служащих и населения.

Особое значение имеет дифференцированный подход. Чем выше риск вмешательства в права человека, тем жестче должны быть требования к применению ИИ. Система распознавания документов может использоваться в упрощенном режиме при соблюдении требований безопасности. Однако система, влияющая на назначение социальной выплаты, определение объекта проверки, миграционный статус или медицинскую помощь, должна проходить предварительную правовую, этическую и техническую экспертизу.

В перспективе такая единая модель может быть обозначена как модель «humanum-sitas» – модель человекоцентричного положения личности в системе цифрового государственного управления. Однако ее значение заключается не в создании новой терминологии как самоцели, а в закреплении принципиального подхода: технология должна быть встроена в правовую систему таким образом, чтобы человек сохранял возможность понимать, контролировать и оспаривать решения, затрагивающие его права.

Заключение

Внедрение технологий искусственного интеллекта в процесс цифровизации государственного управления обладает значительным потенциалом. ИИ способен повысить оперативность предоставления государственных услуг, сократить рутинную нагрузку, выявлять риски, оптимизировать распределение ресурсов и улучшать качество аналитической работы органов публичной власти. Однако положительный эффект не является автоматическим и зависит от того, каким образом технологии включены в систему правовых гарантий.

Научная новизна статьи состоит в систематизации рисков, которые возникают именно при использовании ИИ в публичной сфере: алгоритмической непрозрачности, дискриминации, ошибок данных, чрезмерного вмешательства в частную жизнь, размывания ответственности, деформации доверия, цифрового неравенства, технологической зависимости, киберугроз и неправомерного профилирования. Эти риски связаны с тем, что ИИ способен не только поддерживать, но и фактически предопределять управленческие решения, воздействующие на конституционно-правовой статус человека.

Основной вывод заключается в необходимости разработки единой модели искусственного интеллекта в государственном управлении. Такая модель должна основываться на принципах и механизмах человекоцентричного использования технологий в публичной сфере. Ее целью должно стать не торможение цифровизации, а обеспечение того, чтобы технологическое развитие усиливало, а не ослабляло законность, равенство, справедливость, подотчетность и доверие к государству.

Литература

1. Конституция Российской Федерации: принята всенародным голосованием 12 декабря 1993 г. с изменениями, одобренными в ходе общероссийского голосования 1 июля 2020 г. // Официальный

интернет-портал правовой информации. URL: http://pravo.gov.ru (дата обращения: 19.09.2026).

2. Указ Президента Российской Федерации от 10 октября 2019 г. № 490 «О развитии искусственного интеллекта в Российской Федерации» вместе с Национальной стратегией развития искусственного интеллекта на период до 2030 года (ред. от 15.02.2024) // Официальный интернет-портал правовой информации. URL: http://pravo.gov.ru (дата обращения: 19.09.2026).

3. Федеральный закон от 27 июля 2006 г. № 149-ФЗ «Об информации, информационных технологиях и о защите информации» // Собрание законодательства Российской Федерации. 2006. № 31, ч. 1. Ст. 3448.

4. Федеральный закон от 27 июля 2006 г. № 152-ФЗ «О персональных данных» // Собрание законодательства Российской Федерации. 2006. № 31, ч. 1. Ст. 3451.

5. Федеральный закон от 26 июля 2017 г. № 187-ФЗ «О безопасности критической информационной инфраструктуры Российской Федерации» // Собрание законодательства Российской Федерации. 2017. № 31, ч. I. Ст. 4736.

6. UNESCO. Recommendation on the Ethics of Artificial Intelligence. Paris: UNESCO, 2021. URL: https://www.unesco.org/en/artificial-intelligence/recommendation-ethics (дата обращения: 19.09.2026).

7. Council of Europe. Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law: CETS No. 225. Vilnius, 2024. URL: https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence (дата обращения: 19.09.2026).

8. European Parliament and Council of the European Union. Regulation (EU) 2024/1689 of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) // Official Journal of the European Union. 2024. URL: https://eur-lex.europa.eu/eli/reg/2024/1689/oj (дата обращения: 19.09.2026).

9. Талапина Э. В. Искусственный интеллект и право: новые вызовы для публичного управления // Журнал российского права. 2020. № 9. С. 5–18.

10. Южаков, В. Н. Выявление системных правовых ограничений цифровизации государственного управления: текущее состояние и перспективы / В. Н. Южаков, А. А. Ефремов // Вестник Воронежского государственного университета. Серия: Право. — 2020. — № 3 (42). — С. 110–123.

11. Zuiderwijk A., Chen Y.-C., Salem F. Implications of the Use of Artificial Intelligence in Public Governance: A Systematic Literature Review and a Research Agenda // Government Information Quarterly. 2021. Vol. 38, No. 3. Article 101577.

Ethical and legal risks of introducing artificial intelligence technologies in the process of digitalization of public administration

Panina E.N.

Siberian Law Institute of the Ministry of Internal Affairs of Russia

The article examines the ethical and legal risks of introducing artificial intelligence technologies in the process of digitalization of public administration. The purpose of the work is to identify and systematize the risks arising from the use of intelligent systems in the activities of public authorities. It is proved that the use of artificial intelligence in the public sphere differs from its use in the commercial sector, since algorithmically supported solutions can directly affect the realization of constitutional rights, the public legal status of citizens and trust in state institutions. The risks of algorithmic opacity, discrimination, erroneous data processing, violation of privacy, blurring of public responsibility, digital inequality, technological dependence on developers, cyber threats, unlawful profiling and substitution of individual review by automated assessment are highlighted. Each risk is considered through possible areas of application of AI in public administration: provision of public services, social support, control and supervisory activities, healthcare, migration control, public procurement and law enforcement. It is concluded that it is necessary to develop a unified model of artificial intelligence in the public sphere, including the principles and mechanisms of human-centered use of technology: legality, priority of rights and freedoms, human control, explicitness of decisions, accountability, non-discrimination, data minimization, algorithm audit, and the availability of out-of-court and judicial appeals.

Keywords: artificial intelligence; public administration; digitalization; ethics of artificial intelligence; human rights; algorithmic transparency; digital inequality; public responsibility; human-centered approach; public services.

References

1. The Constitution of the Russian Federation: adopted by popular vote on December 12, 1993 with amendments approved during the all-Russian vote on July 1, 2020 // Official Internet Portal of Legal Information. URL: http://pravo.gov.ru (date of reference: 09/19/2026).
2. Decree of the President of the Russian Federation dated October 10, 2019 No. 490 "On the development of artificial Intelligence in the Russian Federation" together with the National Strategy for the Development of Artificial Intelligence for the period up to 2030 (ed. from 15.02.2024) // Official Internet Portal of Legal Information. URL: http://pravo.gov.ru (date of reference: 09/19/2026).
3. Federal Law No. 149-FZ of July 27, 2006 "On Information, Information Technologies and Information Protection" // Collection of Legislation of the Russian Federation. 2006. No. 31, part 1, Article 3448.
4. Federal Law No. 152-FZ of July 27, 2006 "On Personal Data" // Collection of Legislation of the Russian Federation. 2006. No. 31, Part 1, Article 3451.
5. Federal Law No. 187-FZ of July 26, 2017 "On the Security of the Critical Information Infrastructure of the Russian Federation" // Collection of Legislation of the Russian Federation. 2017. No. 31, ch. I. St. 4736.
6. UNESCO. Recommendation on the ethics of artificial intelligence. Paris: UNESCO, 2021. URL: https://www.unesco.org/en/artificial-intelligence/recommendation-ethics (date of request: 09/19/2026).
7. The Council of Europe. Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law: CETS No. 225. Vilnius, 2024. URL: https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence (date of request: 09/19/2026).
8. The European Parliament and the Council of the European Union. Regulation (EU) 2024/1689 of June 13, 2024, establishing harmonized rules in the field of artificial intelligence (Artificial Intelligence Act) // Official Journal of the European Union. 2024. URL: https://eur-lex.europa.eu/eli/reg/2024/1689/oj (date of access: 09/19/2026).
9. Talapina E. V. Artificial intelligence and law: new challenges for public administration // Journal of Russian Law. 2020. No. 9. pp. 5-18.
10. Yuzhakov, V. N. Identification of systemic legal limitations of digitalization of public administration: current state and prospects / V. N. Yuzhakov, A. A. Efremov // Bulletin of Voronezh State University. Series: Law. — 2020. — № 3 (42). — Pp. 110-123.
11. Zuiderwijk A., Chen Y.-S., Salem F. The consequences of using artificial intelligence in public administration: a systematic review of the literature and a research program // Government Information Quarterly. 2021. Volume 38, No. 3. Article 101577.