

7/22

ПРЕПРИНТЫ

**ЦИФРОВОЕ ОБЩЕСТВО
И ИНФОРМАЦИОННЫЕ
ТЕХНОЛОГИИ. ИННОВАЦИИ
DIGITAL SOCIETY AND
INFORMATION TECHNOLOGIES
INNOVATIONS**

**ЦИФРОВОЕ ОБЩЕСТВО
И ИНФОРМАЦИОННЫЕ
ТЕХНОЛОГИИ. ИННОВАЦИИ
DIGITAL SOCIETY AND
INFORMATION TECHNOLOGIES
INNOVATIONS**

В. Н. Южаков, Э. В. Талапина, И. А. Черешнева

**АНАЛИЗ ПРАВОВЫХ СПОСОБОВ
ПРЕДОТВРАЩЕНИЯ НАРУШЕНИЙ
ФУНДАМЕНТАЛЬНЫХ ПРАВ ЧЕЛОВЕКА
ПРИ ИСПОЛЬЗОВАНИИ ИСКУССТВЕННОГО
ИНТЕЛЛЕКТА В ГОСУДАРСТВЕННОМ УПРАВЛЕНИИ**

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ
УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА
И ГОСУДАРСТВЕННОЙ СЛУЖБЫ
ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ»
(РАНХиГС)

ПРЕПРИНТ
(НАУЧНЫЙ ДОКЛАД)

по теме:

**АНАЛИЗ ПРАВОВЫХ СПОСОБОВ ПРЕДОТВРАЩЕНИЯ НАРУШЕНИЙ
ФУНДАМЕНТАЛЬНЫХ ПРАВ ЧЕЛОВЕКА ПРИ ИСПОЛЬЗОВАНИИ
ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ГОСУДАРСТВЕННОМ
УПРАВЛЕНИИ**

Южаков В.Н., директор ЦТГУ, д.ф.н., проф., 0000-0002-5687-1863,

yuzhakov-vn@ranepa.ru

Талапина Э.В., в.н.с. ЦТГУ, д.ю.н., 0000-0003-3395-3126, talapina-ev@ranepa.ru

Черешнева И.А., н.с. ЦТГУ, 0000-0001-8135-4166, chereshneva-ia@ranepa.ru

Москва 2022

Аннотация

В настоящий момент ни одна страна в мире не выработала целостного нормативного подхода к регулированию применения искусственного интеллекта. Между тем, в цифровых условиях риски нарушений прав человека возрастают, что **актуализирует** необходимость одновременной разработки правовой защиты прав человека при освоении государственным управлением возможностей технологий искусственного интеллекта. В этой связи **целями** исследования стали анализ текущего состояния правового регулирования вопросов освоения ИИ государственным управлением и подготовка предложений по правовому обеспечению предотвращения и преодоления рисков нарушений прав человека при использовании ИИ в государственном управлении. В ходе исследования проанализированы научные публикации по данной тематике, нормативные правовые акты международного и национального уровня, в том числе зарубежных государств. В препринте представлены **основные результаты** исследования: итоги анализа состояния правового регулирования вопросов освоения ИИ в государственном управлении; итоги анализа потенциальных рисков нарушений прав человека при использовании ИИ в государственном управлении; авторская систематизация рисков нарушений прав человека при использовании ИИ в государственном управлении, в том числе с учетом изменений институциональной системы защиты прав; предложения по правовым способам предотвращения и преодоления рисков нарушений прав человека при использовании ИИ в государственном управлении. Исследование позволило сделать **выводы**: о недостаточном внимании к аспекту прав человека при прогнозировании, регулировании и использовании ИИ в государственном управлении, о необходимости проработать систему правовых способов предотвращения и преодоления рисков нарушений прав человека при использовании ИИ в государственном управлении. Неурегулированность процесса использования искусственного интеллекта в государственном управлении, влекущая риски нарушений прав человека, определяет **научную новизну** и потенциальную **практическую значимость** предлагаемого вниманию читателя исследования. Авторский подход предполагает продвижение критерия человекоцентричности при применении технологий искусственного интеллекта в государственном управлении. Полученные результаты и выводы могут быть учтены и использованы для разработки и реализации государственной политики в области обеспечения прав человека при использовании ИИ в российском государственном управлении.

Ключевые слова:

Права человека, государственное управление, данные, искусственный интеллект, персональные данные, обработка больших данных, алгоритм

Коды JEL Classification

H11; H83; K38

RUSSIAN PRESIDENTIAL ACADEMY OF NATIONAL ECONOMY
AND PUBLIC ADMINISTRATION (RANEPА)

PREPRINT
(SCIENTIFIC REPORT)

**THE ANALYSIS OF LEGAL MEANS TO PREVENT FUNDAMENTAL HUMAN
RIGHTS VIOLATIONS DUE TO THE USE OF ARTIFICIAL INTELLIGENCE
IN PUBLIC ADMINISTRATION**

Yuzhakov Vladimir N., director, Center of Public Administration Technologies, Dr. Sci. (Philosophy), professor, ORCID 0000-0002-5687-1863, yuzhakov-vn@ranepa.ru

Talapina Elvira V., lead researcher, Center of Public Administration Technologies, Dr. Sci. (Law), ORCID 0000-0003-3395-3126, talapina-ev@ranepa.ru

Chereshneva Irina A. researcher, Center of Public Administration Technologies,
ORCID 0000-0001-8135-4166, chereshneva-ia@ranepa.ru

Abstract

Nowadays none country in the world has developed a holistic regulatory approach to regulation of artificial intelligence use. Meanwhile, in the digital landscape the risks of human rights violations are increasing, which is **gained newfound relevance** of the need to simultaneous development of legal protection of human rights in mastery of technology and artificial intelligence capabilities by public administration. In this regard **the objectives** of this paper are to analyze the current state of legal regulation of AI use in public administration and to formulate of proposals for legal ground-work for preventing and overcoming the risks of human rights violations due to the use of AI in public administration. As a part of the study scientific publications on this topic, international and national laws and regulations, including foreign countries have been analyzed. **The main results** of the preprint are the results of the analysis of the state of legal regulation of AI use in public administration; the results of the analysis of potential risks of human rights violations due to the use of AI in public administration; the author's systematization of the risks of human rights violations due to the use of AI in public administration, as varied the institutional system of rights protection; the proposals for legal means to prevent and overcome the risks of human rights violations due to the use of AI in public administration. The study allows for **the conclusion** about the lack of attention to the human rights aspect in forecasting, regulating and using of AI in public administration as well as the necessity to design a system of legal means to prevent and overcome the risks of human rights violations due to the use of AI in public administration. The lack of proper regulation of AI use in public administration, which involves the risks of human rights violations, determines **the scientific novelty and potential practical significance** of the study brought to attention of readers. The author's concept involves the promotion of a human-centered approach to the use of artificial intelligence technology in public administration. Its findings and results can be taken into account and used for the formulation and state policy implementation in ensuring human rights due to the use of AI in the Russian public administration.

Key words:

Human rights, public administration, data, artificial intelligence, personal data, big data processing, algorithm

JEL Classification

H11; H83; K38

Оглавление

Введение	6
1. Зарубежный и российский опыт правового регулирования вопросов освоения ИИ в государственном управлении	7
2. Потенциальные риски нарушений прав человека при использовании ИИ в государственном управлении	11
3. Систематизация рисков нарушений прав человека при использовании ИИ в государственном управлении	15
3.1. Анализ подходов к систематизации рисков нарушений прав человека при использовании ИИ в государственном управлении	15
3.2. Предложения по систематизации рисков нарушений прав человека при использовании ИИ в государственном управлении	18
4. Правовые способы преодоления рисков нарушений прав человека при использовании ИИ в государственном управлении	28
Заключение	31
Благодарности	33
Список источников	34

Введение

На настоящий момент еще ни одна страна в мире не выработала целостного нормативного подхода к регулированию применения искусственного интеллекта. Между тем, разработка такого регулирования должна учитывать и риски нарушения прав человека, возникающие при освоении ИИ государственным управлением. Нельзя допустить, чтобы человек остался без действенной системы защиты своих прав.

Целью настоящего исследования является анализ текущего состояния правового регулирования вопросов освоения ИИ в государственном управлении и подготовка предложений по правовому обеспечению предотвращения и преодоления рисков нарушений прав человека при использовании ИИ в государственном управлении.

Для достижения цели выполнены следующие задачи:

- 1) анализ текущего состояния правового регулирования вопросов освоения ИИ в государственном управлении;
- 2) выявление потенциальных рисков нарушений прав человека при использовании ИИ в государственном управлении;
- 3) систематизация рисков нарушений прав человека при использовании ИИ в государственном управлении, в том числе с учетом изменений институциональной системы защиты прав;
- 4) подготовка предложений по правовым способам предотвращения и преодоления рисков нарушений прав человека при использовании ИИ в государственном управлении.

Новизна такого исследования определяется неурегулированностью процесса использования искусственного интеллекта в государственном управлении, что влечет риски нарушений прав человека. При этом акцент на права человека несвойственен деятельности в сфере государственного управления, поскольку реформирование нацелено на государственный аппарат как таковой. Авторский подход предполагает продвижение критерия человекоцентричности при применении технологий искусственного интеллекта в государственном управлении.

Полученные результаты могут быть использованы для разработки соответствующей государственной политики в области обеспечения прав человека при использовании ИИ в российском государственном управлении.

1. Зарубежный и российский опыт правового регулирования вопросов освоения ИИ в государственном управлении

Если исходить из того тезиса, что «обузданная» технология искусственного интеллекта может обеспечить технологическое лидерство государства, то для анализа зарубежного опыта ее использования в сфере государственного управления целесообразно остановиться на нормативной составляющей тех стран, которые в настоящее время находятся в числе пионеров, – США, Китай, Европейский Союз.

Соединенные Штаты Америки

В части правового регулирования ИИ США воздерживаются от введения горизонтального законодательства в цифровой сфере, вместо этого сосредотачивая внимание на отраслевых законах и инновациях частного сектора, особенно в отношении технологических гигантов (например, Google, Amazon, Facebook и др.) и ведущих университетов (например, MIT). Это скорее рекомендательный подход, т.е. предоставление компаниям, инвестирующим в исследовательские проекты, юридических рекомендаций, а также устранение барьеров на пути развития инноваций.

Национальная инициатива по искусственному интеллекту (The National Artificial Intelligence Initiative) была учреждена Законом о Национальной инициативе по искусственному интеллекту 2020 года [1], принятие которого поддержали обе партии Конгресса США 1 января 2021 года в рамках Закона о бюджетных ассигнованиях на национальную оборону (принимается на каждый финансовый год) на 2021 финансовый год. В поддержку Инициативы Закон о Национальной инициативе по искусственному интеллекту наделяет Президента США, действующего через аппарат Национальной стратегии США в области искусственного интеллекта, межведомственный комитет (Специальный комитет по ИИ) и глав федеральных органов государственной власти, рядом полномочий, в том числе: поддерживать программы обучения ИИ и подготовки кадров; осуществлять планирование и координацию федеральной межведомственной деятельности в области искусственного интеллекта; поддерживать сети междисциплинарных научно-исследовательских институтов в области искусственного интеллекта и др.

Регулирование использования ИИ федеральным правительством осуществляется в рамках AI in Government Act of 2020, который выступает составной частью Сводного закона об ассигнованиях 2021 года (Consolidated Appropriations Act) [2]. Указ Президента США от 3 декабря 2020 года № 13960 «О содействии использованию надежного искусственного интеллекта федеральными органами власти» устанавливает принципы использования ИИ в

федеральных органах власти (законность и уважение ценностей американской нации; целенаправленность и ориентация на результат; точность, надежность и эффективность; безопасность и надежность; понятность; ответственность и прослеживаемость; наблюдение и контроль; транспарентность; подотчетность) [3], устанавливает общую политику по внедрению принципов, предписывает федеральным органам описать и включить в перечень свои варианты использования ИИ, а также призывает Администрацию общих служб США и Службу управления персоналом расширить опыт внедрения ИИ в федеральных органах [4].

Несмотря на объявление о разработке общих стандартов надежного ИИ, перечисленные принципы достаточно широки. Это позволяет каждому федеральному органу власти разработать такое нормативное регулирование, которое будет нацелено на конкретную отрасль. При этом предполагается, что в будущем Администрация Президента Байдена представит новый «билль о правах» для того, чтобы ограничить потенциальный вред, который может причинить технология искусственного интеллекта [5]. Несмотря на это, в части регулирования ИИ США все также будут руководствоваться подходом, ориентированным на потребности рынка, тем самым, избегая чрезмерного регулирования [6].

Китай

В 2015 году китайское правительство выдвинуло национальный стратегический план «Сделано в Китае – 2025» (Made in China 2025), который представляет собой инициативу, направленную на закрепление за Китаем позиции мирового лидера в высокотехнологичных отраслях. Цель данной инициативы заключается в уменьшении зависимости КНР от импорта иностранных технологий и инвестировании значительных средств в собственные инновации для создания китайских компаний, способных конкурировать как внутри страны, так и на глобальном уровне [7]. Наряду с вышеупомянутым планом в 2017 году была принята Программа развития ИИ нового поколения (Next Generation AI Development Plan). Оба этих документа ставят перед Китаем весьма амбициозную цель – стать мировым лидером в области ИИ к 2030 году.

Программой развития ИИ нового поколения определяются стратегические цели, со сроками их достижения, а также описываются бюрократические механизмы работы и источники финансирования [8]. При этом нормативных актов, регламентирующих использование ИИ, нет.

Европейский Союз

Европейская комиссия начала свою работу по регулированию ИИ в 2018 году, опубликовав Европейскую стратегию ИИ, а также создав Группу экспертов высокого уровня и представив Скоординированный план в области ИИ (Coordinated Plan on Artificial

Intelligence) [9] в целях оказания содействия в реализации европейского подхода к ИИ – AI made in Europe («ИИ, сделанный в Европе»). Главный итог - Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Проект Регламента ЕС о гармонизированных правилах в отношении ИИ (Artificial Intelligence Act)) [10].

Цель данного регламента состоит в улучшении функционирования внутреннего рынка путем создания единой правовой базы, в том числе для разработки, маркетинга и использования искусственного интеллекта сообразно ценностям ЕС. Круг применения технологии ИИ предлагается ограничить: в качестве примера можно привести запрет на внедрение систем социального скоринга для государственных нужд. Кроме того, в документе определяются высокорисковые области применения ИИ и требования к процессам, системам и данным, используемым в таких областях. Если применение ИИ можно отнести к области ограниченного или минимального риска, то в первом случае предусматривается ограниченное регулирование, которое охватывает вопросы прозрачности взаимодействия с системой ИИ, а второе – отсутствие регулирования. Соответственно для областей высокого риска предусматривается процедура подтверждения соответствия.

В целом проект Европейского регламента по ИИ выстроен на базовой идее соблюдения и эффективного применения существующего законодательства об основных правах. Отметим, что проектом Европейского регламента по ИИ также предусматривается создание Европейского совета по ИИ, регуляторных песочниц и введение добровольных кодексов поведения, однако только для тех областей применения ИИ, которые не относятся к высокорисковым [10].

Таким образом, для ЕС характерен регулирующий подход, центральное место в котором отводится этическим соображениям, соответствующим основным ценностям в области прав человека и демократическим принципам. Тем самым, Еврокомиссия надеется достичь еще одного «брюссельского эффекта», подобного GDPR.

Каждый из проанализированных подходов (американский, китайский, европейский) отражает устойчивые черты их сформированных правовых систем, с их доцифровыми характеристиками. Очевидно, учет подобных национальных особенностей должен быть произведен и в России, поскольку полного заимствования быть не может.

Необходимо отметить, что опыт правового регулирования использования информационных технологий в российском государственном управлении довольно обширен и насчитывает десятилетия. В частности, речь идет о создании еще в Советском Союзе автоматизированных систем управления (АСУ) [11], а затем и создании и использовании государственных информационных систем (ГИС). В этом смысле проблемы нарушений прав

человека при использовании новых технологий в государственном управлении имеют собственную историю, положительный и негативный опыт, который необходимо учитывать.

Тем не менее, развитие законодательства в области регулирования ИИ – вопрос первой необходимости. Так, в дорожной карте развития «сквозной» цифровой технологии «Нейротехнологии и искусственный интеллект» [12] в качестве ограничений в области «Нормативное регулирование» назывались:

- отсутствие полноценной системы нормативно-технического регулирования в области ИИ;
- сложность административно-правового и нормативно-технического порядка тестирования и внедрения разработок в сфере ИИ;
- несовместимость некоторых положений законодательства о защите данных с технологиями ИИ.

Следует отметить, что российский опыт правового регулирования использования ИИ находится на начальной стадии, что, впрочем, соответствует актуальному уровню большинства государств. Несмотря на представленный выше опыт объединенной Европы, США и Китая (больше программного характера), приходится констатировать, что в настоящее время в мире нет ни одной страны (или наднациональной организации), законодательство которой явно учитывало бы проблемные характеристики алгоритмических систем в целом. За некоторыми исключениями, также не существует специальных правил для систем ИИ и интеллектуальной робототехники в частности.

2. Потенциальные риски нарушений прав человека при использовании ИИ в государственном управлении

Права человека воспринимаются как фундаментальные, уникальные, основные, защищающие человека как человека. Именно это обстоятельство делает их совершенно особенными в любой юридической системе, независимо от того, на чем эта система основывается: признание естественных прав, присущих человеку, теологическая вера в то, что уникальные права человека – от Бога, идея Хабермаса о связи прав человека и демократии, или что-то еще.

Прежде всего, необходимо разобраться, в отношении каких прав человека использование новых технологий способно вызвать наиболее существенные и очевидные риски. Для информационного общества в нашу цифровую эпоху некоторые права, регулируемые Хартией ЕС об основных правах, особенно актуальны, в частности, право на уважение частной жизни (ст. 7), защиту персональных данных (ст. 8), право на информацию (ст. 11), право на собственность (ст. 17), в частности, авторское право, которое может быть нарушено цифровыми «инструментами», свобода ведения бизнеса (ст. 16). Права на эффективное средство правовой защиты и беспристрастное судебное разбирательство (ст. 47) также являются важными. Эти судебные права регулируются в качестве основных прав. Запрет на дискриминацию (ст. 21) тоже является частой проблемой в цифровую эпоху. Проблемы дискриминации часто возникают между частными субъектами в онлайн-общении (возраст, пол, религия и т.д.), и, таким образом, отдельно высвечивается проблема прямого действия прав человека в отношениях между частными лицами. Все эти права влияют на развитие, взаимоотношения, права и обязанности в цифровом обществе.

Применение ИИ и его алгоритмов в государственном управлении довольно сложно освободить от рисков принятия автоматизированных решений, а именно:

а) Стандартизация поведения.

С распространением автоматизированных решений возникает риск стандартизации, а именно риск того, что индивидуальное поведение большого количества людей будет в результате выровнено и станет более единообразным [13]. С одной стороны, это может быть связано с тем, что с помощью автоматизированной системы операторы системы (или автономно сама система) могут устанавливать стандарты, которые затем могут единообразно и эффективно применяться к различным ситуациям.

Люди склонны согласовывать свое собственное поведение с поведением других людей. Для отдельного человека одобрение со стороны масс может сделать вариант решения

заслуживающим доверия, но также может создать препятствие для того, чтобы не отклоняться от него. В зависимости от дизайна системы ИИ, может наблюдаться концентрация моделей поведения и сближение индивидуумов. Сумма вариантов, доступных для индивидуума, может иметь тенденцию к сокращению и сосредоточению на основном поведении и решениях. В итоге реализация индивидуальности может потребовать больших усилий и больших затрат и даже привести к социальному разделению.

Представляется, что ущемление автономии личности с помощью автоматизированных систем не должно приводить к тому, чтобы контроль со стороны третьих лиц заменял самоопределение, защищенное основными правами. В строго либеральном смысле это также может включать в себя свободу совершать ошибки. Более того, перспектива «совершенного» правового государства, в смысле все более эффективного и, возможно, даже повсеместного применения закона, является не совсем привлекательной перспективой. Поэтому наличие технологий для реализации совершенного правоприменения заставляет нас решать, где и когда общество может захотеть сохранить свободу не подчиняться правилу (правилам) закона.

Таким образом, главными последствиями реализации иска стандартизации поведения граждан могут стать чрезмерная управляемость общества, а также социальное разделение (фрагментация) общества.

б) Недостаточная понятность.

Автоматизированные системы могут определять характеристики, склонности, цели и намерения людей в ранее неизвестной глубине и детализации и, таким образом, вычислять прогнозы относительно их будущего поведения. Человеческие когнитивные способности не могут угнаться за этим, так что человеческая способность реально постичь конкретные процессы принятия решений автоматизированных систем достигает своего предела. Если автоматизированная система определяет определенный контекст и основывает на нем свое решение, существует опасность, что автоматизированная процедура больше не будет понятна человеку.

Возможности индивидуального самоопределения ущемляются, если человек не знает, какие критерии используются автоматизированной системой. К примеру, в немецком законодательстве индивидуальное самоопределение является выражением общего права на развитие личности в соответствии со статьей 2(1), 1(1) Основного закона. Свободное развитие личности и независимый образ жизни защищены, лишь если принимается во внимание индивидуальность каждого.

Последствиями реализации риска недостаточной понятности может стать незнание граждан о нарушении их прав и, как следствие, отсутствие их защиты (особенно в рамках тотального наблюдения, о котором человеку неизвестно).

в) Информационная асимметрия.

Использование автоматизированных систем может привести к информационной асимметрии, в которой преимущество одного лица возникает именно благодаря информации о и/или против другого лица. Использование автоматизированных систем дает возможность оценить, какие индивидуальные потребности и цели есть или могут быть обеспечены, независимо от того, осознает ли само лицо свои собственные потребности и цели [13].

Реализация риска информационной асимметрии приводит к дискриминации, а также навязыванию услуг (в случае с частными субъектами – коммерческих услуг, а в случае с государством – публичных). Зачастую невозможность реализовать свои права иначе чем электронным способом, ведет к навязыванию цифровых сервисов (когда граждане вынужденно используют технологии, поскольку нет альтернативы, или не используют их вообще, усугубляя цифровое неравенство).

Традиционный способ отгородиться от вездесущей слежки – это сослаться на «1984» или, в наши дни, «систему социального кредита», т.е. вызвать угрозу оруэлловского или китайского государства наблюдения, которое может злоупотребить персональными данными для подавления индивидуализма. На другой чаше весов – умное правоприменение. Умное правоприменение определяется как автоматизированное обнаружение подозрительного поведения. Поведение является подозрительным в этом смысле, если оно указывает, с достаточной точностью и вероятностью, что некоторая форма несоблюдения правовых норм имела место, происходит или произойдет в будущем [14].

Правовые меры реагирования на описанные здесь риски можно разделить на инструменты, имеющие прямое и косвенное воздействие. Так, реагирование на стандартизацию поведения обсуждается в рамках оправданности и подотчетности «подталкивания» (system nudges – система подталкивания). Например, внимание может быть сосредоточено на том, принесет ли побуждаемое решение больше пользы, руководствуются ли заинтересованные лица лучшими решениями на основе собственных стандартов, или же их первоначальное поведение является предвзятым и ошибочным.

В юридическом плане особое значение имеет проблема *юридической значимости влияния на людей*. Основной вопрос здесь заключается в том, чтобы определить, когда такой потенциал влияния является юридически значимым. Когда, следовательно, правовая система должна рассматривать его как риск для самоопределения личности? В принципе, интенсивность потенциала влияния может быть определена только самим индивидом, то есть

индивидуально. Воспринимаемое давление с целью изменения поведения в значительной степени зависит от индивидуального опыта и вряд ли может быть типизировано. Чем больше персональных данных используют автоматизированные системы для влияния на поведение, тем менее прозрачными они кажутся, и тем больше они влияют на бессознательные и иррациональные когнитивные или преднамеренные процессы человека. Использование случайно появляющихся критериев может оправдать запрет на автоматизированное влияние на индивидуальное самоопределение (использование критериев, которые не являются предсказуемыми и понятными в соответствии с текущим горизонтом ожиданий человека).

3. Систематизация рисков нарушений прав человека при использовании ИИ в государственном управлении

3.1. Анализ подходов к систематизации рисков нарушений прав человека при использовании ИИ в государственном управлении

В контексте обеспечения защиты прав человека при использовании ИИ вообще и в государственном управлении в частности *в отечественной юридической литературе* выделяются следующие ключевые риски, имеющие непосредственное юридическое значение.

1. Риск отсутствия прозрачности сложных систем ИИ, который может быть выражен формулой «увеличение сложности применяемых алгоритмов обратно пропорционально степени их прозрачности и возможностям человека по их пониманию и объяснению» [15]. Тем самым, отсутствие прозрачности может привести к осложнению процесса оспаривания со стороны граждан судебных и иных административных решений, в процессе принятия которых были использованы системы ИИ. Этот риск также «покрывает» «низкую способность алгоритмов к выделению и пониманию причинно-следственных связей» [15], ввиду того, что жизнь гораздо сложнее и многограннее простой (или даже сложной) последовательности шагов, а все многообразие ситуаций записать в двоичный код невозможно. Кроме того, нельзя забывать, что зачастую люди принимают решения, основываясь не только на логике, но и руководствуясь эмоциями, импульсивно, иррационально. Тем самым, система ИИ может «просчитать» тот или иной вариант как нерациональный и, следовательно, маловероятный, тогда как в реальности такой вариант будет иметь место.

2. Риск предвзятости алгоритмов напрямую вытекает как из потенциальной предвзятости со стороны разработчиков алгоритмов, так и из нашего общества *per se*, на основе достижений и недостатков которого система ИИ будет осуществлять свое обучение. Например, в деле *State v. Loomis* была проведена алгоритмическая оценка риска совершения лицом рецидива преступления. Несмотря на то, что судом был вынесен обвинительный приговор, а доводы обвиняемого не были приняты во внимание, в качестве одной из обязательных рекомендаций нижестоящим судам была дана рекомендация относительно того, что «при оценке осуществляется непропорциональная классификация правонарушителей из числа меньшинств как имеющих более высокий риск совершения» [16], т.е. фактически речь шла о расовой дискриминации.

3. Риск уязвимости к ошибкам (статистическим, вычислительным, субъективной природы) [15] включает в себя не только «простую» (в том смысле, как и любой программный код) подверженность систем ИИ различного рода ошибкам, но и «способность ИИ к

"самоусилению" исходной информации, на основе которой делаются все последующие выводы» [15], что, в свою очередь, может еще более усилить риск предвзятости алгоритмов.

4. Риск регуляторного потенциала систем ИИ связывают с тем обстоятельством, что, как и право, «алгоритмы нацелены на управление будущими процессами при помощи математических моделей, по структуре напоминающих нормы права («если... то... иначе»)» [15], однако в отличие от права, которое характеризуется открытостью и гибкостью в правоприменении, алгоритмы, являясь «закрытыми системами, чужды учету сопутствующих обстоятельств, на которые они не запрограммированы, что создает существенные риски "замораживания" тех или иных социальных отношений в заданных конфигурациях» [15]. Сказанное, в свою очередь, может привести к риску выпадания из орбиты регулирующего воздействия таких общечеловеческих ценностей, как справедливость, нравственность, честь, достоинство, добросовестность и др.

В зарубежных исследованиях проводится более линейная связь между правовыми рисками и правами человека (таблица 1).

Таблица 1

Правовые риски, связанные с ИИ, и затронутые ими права человека

Правовые риски, связанные с ИИ	Принципы прав человека, которые могут быть нарушены
Недостаток алгоритмической транспарентности	Справедливое судебное разбирательство и надлежащая правовая процедура; эффективные средства правовой защиты; социальные права и доступ к государственным услугам; права на свободные выборы
Уязвимости в области кибербезопасности	Право на неприкосновенность частной жизни; свобода выражения мнений и свободный поток информации
Несправедливость, предвзятость и дискриминация	Устранение всех форм дискриминации в отношении женщин; равные права мужчин и женщин; осуществление прав детей без дискриминации; равенство перед законом, равная защита со стороны закона без дискриминации; осуществление установленных прав без дискриминации; недискриминация, право на жизнь трудовых мигрантов; право на свободу и безопасность личности; запрет на дискриминацию по признаку ограниченных возможностей; право на справедливое судебное разбирательство; право на свободу от дискриминации
Недостаток оспоримости	Право на эффективные средства правовой защиты; доступ к правосудию
Правосубъектность, субъектность, нравственность	Право на всеобщее признание в качестве физического лица перед законом; право на равенство; устранение всех форм дискриминации
Проблемы в сфере интеллектуальной собственности	Право на единоличное или совместное владение собственностью; право свободно участвовать в культурной жизни общества, наслаждаться искусством, участвовать в научном прогрессе и пользоваться его благами; право на защиту его моральных и материальных интересов, являющихся результатом научных, литературных или художественных трудов, автором которых он является
Неблагоприятное воздействие на работников	Право на социальное обеспечение; запрет дискриминации в отношении осуществления прав на труд, на свободный выбор работы, на справедливые и благоприятные условия труда, на защиту от безработицы, на равную оплату за равный труд, на справедливое и удовлетворительное воз-

Правовые риски, связанные с ИИ	Принципы прав человека, которые могут быть нарушены
	награждение; право на труд, включая право каждого на возможность зарабатывать себе на жизнь трудом, который он/она свободно выбирают или принимают; право лиц с ограниченными возможностями на труд наравне с другими людьми
Проблемы в области неприкосновенности частной жизни и защиты данных	Право мигранта на неприкосновенность частной жизни; уважение частной жизни лиц с ограниченными возможностями; право на уважение частной и семейной жизни; право на неприкосновенность частной жизни и защиту данных; неприкосновенность частной жизни детей; защита неприкосновенности пожилых людей, их частной и личной жизни
Проблема ответственности за причиненный вред	Право на жизнь, право на эффективные средства правовой защиты
Отсутствие ответственности за причиненный ущерб	Право на жизнь, право на эффективные средства правовой защиты [17]

3.2. Предложения по систематизации рисков нарушений прав человека при использовании ИИ в государственном управлении

В целях создания методологических основ для оценки ИИ, используемого в государственном управлении, на наличие рисков основным правам человека целесообразно прибегнуть к матрице, которая позволит воплотить принцип объективной экспертизы, а именно: 1) проверить каждое право человека на наличие предполагаемых рисков его нарушения и 2) проверить каждый риск нарушения прав человека на его наличие при осуществлении каждой функции государственного управления. Одновременно анализ предполагает типовой перечень способов преодоления рисков, что можно охарактеризовать также как способ исправления «угрожающей» ситуации.

Проведенный нами анализ российского и зарубежного опыта позволяет выдвинуть следующие предложения по систематизации рисков нарушений прав человека при использовании ИИ в государственном управлении (*таблица 2*), а также провести систематизацию рисков нарушений прав человека в отношении каждой функции государственного управления (*таблица 3*).

Таблица 2.

Систематизация рисков нарушений прав человека при использовании ИИ в государственном управлении

Право человека	Риски	Способы преодоления рисков	Применение в целом и в сфере государственного управления
Право на жизнь	Недостаточная понятность функционирования ИИ Риск предвзятости алгоритмов Риск уязвимости к ошибкам	Объяснимый ИИ; Право на объяснение; Повышение информированности граждан; Подотчетность при использовании алгоритмов в процессе принятия решений в государственном секторе; Надзор со стороны контролирующих органов и юридическая ответственность; Оценка алгоритмического воздействия	Судебные решения по уголовным делам
Право на достоинство личности	Риск стандартизации поведения граждан Дискриминация Риск предвзятости алгоритмов Риск уязвимости к ошибкам	Объяснимый ИИ; Право на объяснение	Оказание государственных услуг
Право на свободу и личную неприкосновенность	Риск стандартизации поведения граждан Дискриминация Информационная асимметрия Риск предвзятости алгоритмов Риск уязвимости к ошибкам	Подотчетность при использовании алгоритмов в процессе принятия решений в государственном секторе; Информационное самоопределение; Гарантии защиты персональных данных	Судебные решения по уголовным делам

Продолжение таблицы 2

Право человека	Риски	Способы преодоления рисков	Применение в целом и в сфере государственного управления
<p>Право на неприкосновенность частной жизни, личную и семейную тайну, защиту своей чести и доброго имени</p>	<p>Риск стандартизации поведения граждан Дискриминация Риск предвзятости алгоритмов Риск уязвимости к ошибкам</p>	<p>Право на объяснение; Информационное самоопределение; Нормативно закрепляемые требования к обороту данных: данные должны быть ограничены минимальным объемом (1); данные должны храниться только в течение ограниченного времени (2); данные должны быть анонимными, где это возможно. Подотчетность при использовании алгоритмов в процессе принятия решений в государственном секторе; Гарантии защиты персональных данных; Проектируемая конфиденциальность; Осуществление надзора независимым органом по защите данных; Человек в контуре управления (humans-in-the-loop) – участие человека в цикле обработки данных</p>	<p>Оказание государственных услуг; Контрольно-разрешительная деятельность</p>
<p>Право на тайну переписки, телефонных переговоров, почтовых, телеграфных и иных сообщений</p>	<p>Недостаточная понятность функционирования ИИ; Дискриминация</p>	<p>Подотчетность при использовании алгоритмов в процессе принятия решений в государственном секторе; Проектируемая конфиденциальность</p>	<p>Разрешительная деятельность Уголовный процесс Правоохранительная деятельность</p>
<p>Право на информацию</p>	<p>Информационная асимметрия; Цифровое неравенство Дискриминация Риск предвзятости алгоритмов Риск уязвимости к ошибкам</p>	<p>Объяснимый ИИ; Право на объяснение; Информационное самоопределение; Право на доступ к Интернету; Право на забвение; Подотчетность при использовании алгоритмов в процессе принятия решений в государственном секторе</p>	<p>Оказание государственных услуг Позитивное государственное управление Нормотворческая деятельность</p>

Продолжение таблицы 2

Право человека	Риски	Способы преодоления рисков	Применение в целом и в сфере государственного управления
Право на участие в управлении делами государства	Информационная асимметрия; Цифровое неравенство Дискриминация Риск предвзятости алгоритмов Риск уязвимости к ошибкам Недостаточная понятность функционирования ИИ	Право на объяснение; Информационное самоопределение; Право на доступ к Интернету; Объяснимый ИИ; Подотчетность при использовании алгоритмов в процессе принятия решений в государственном секторе; Человек в контуре управления (humans-in-the-loop) – участие человека в цикле обработки данных; Проектируемая оспоримость	Нормотворческая деятельность Юридическая экспертиза
Право избирать и быть избранными в органы государственной власти и органы местного самоуправления, а также участвовать в референдуме	Дискриминация Цифровое неравенство Риск предвзятости алгоритмов Риск стандартизации поведения граждан Информационная асимметрия	Право на объяснение; Информационное самоопределение; Право на доступ к Интернету	Позитивное государственное управление
Право на участие в отправлении правосудия	Дискриминация Цифровое неравенство Риск предвзятости алгоритмов Риск стандартизации поведения граждан Информационная асимметрия Риск излишней самонадеянности в отношении автоматизации	Информационное самоопределение; Подотчетность при использовании алгоритмов в процессе принятия решений в государственном секторе	Отправление правосудия
Право на объединение	Риск стандартизации поведения граждан Информационная асимметрия	Информационное самоопределение; Право на доступ к Интернету	Позитивное государственное управление
Право на собрание, митинги и шествия	Риск стандартизации поведения граждан Информационная асимметрия Цифровое неравенство	Информационное самоопределение	Позитивное государственное управление

Продолжение таблицы 2

Право человека	Риски	Способы преодоления рисков	Применение в целом и в сфере государственного управления
Право частной собственности	Цифровое неравенство Риск предвзятости алгоритмов Риск стандартизации поведения граждан	Информационное самоопределение; Право на забвение	Оказание государственных услуг Позитивное государственное управление
Право на свободное использование своих способностей и имущества для предпринимательской и иной не запрещенной законом экономической деятельности	Дискриминация Цифровое неравенство Риск предвзятости алгоритмов Риск стандартизации поведения граждан	Информационное самоопределение; Право на забвение; Человек в контуре управления (humans-in-the-loop) – участие человека в цикле обработки данных; Право на объяснение	Оказание государственных услуг (включая аутсорсинг административно-управленческих процессов)
Право на землю	Цифровое неравенство Риск предвзятости алгоритмов Риск стандартизации поведения граждан Информационная асимметрия	Проведение регулярных оценок на предмет репрезентативности наборов данных, учет наличия или отсутствия в наборах данных элементов предвзятости; Право на объяснение	Государственная поддержка уязвимых категорий граждан
Право на жилище	Цифровое неравенство Риск предвзятости алгоритмов Риск стандартизации поведения граждан Информационная асимметрия	Проведение регулярных оценок на предмет репрезентативности наборов данных, учет наличия или отсутствия в наборах данных элементов предвзятости; Право на объяснение	Государственная поддержка уязвимых категорий граждан
Право каждого свободно распоряжаться своими способностями к труду, выбирать род деятельности и профессию	Дискриминация Цифровое неравенство Риск предвзятости алгоритмов Риск стандартизации поведения граждан Информационная асимметрия	Право на забвение; Информационное самоопределение; Человек в контуре управления (humans-in-the-loop) – участие человека в цикле обработки данных	Нормотворческая деятельность Государственная служба Позитивное государственное управление
Право на труд в условиях, отвечающих требованиям безопасности и гигиены	Дискриминация Цифровое неравенство Риск предвзятости алгоритмов Риск стандартизации поведения граждан	Человек в контуре управления (humans-in-the-loop) – участие человека в цикле обработки данных	Государственная служба/Позитивное государственное управление

Продолжение таблицы 2

Право человека	Риски	Способы преодоления рисков	Применение в целом и в сфере государственного управления
Право на вознаграждение за труд без какой бы то ни было дискриминации и не ниже установленного федеральным законом минимального размера оплаты труда	Дискриминация Цифровое неравенство Риск предвзятости алгоритмов Риск стандартизации поведения граждан	Изменение системы социального обеспечения; Право на объяснение	Позитивное государственное управление
Право на защиту от безработицы	Дискриминация Цифровое неравенство Риск предвзятости алгоритмов Риск стандартизации поведения граждан	Проведение переподготовки кадров; Переориентация и адаптация системы образования; Изменение системы социального обеспечения	Оказание государственных услуг
Право на отдых	Дискриминация Цифровое неравенство Риск предвзятости алгоритмов Риск стандартизации поведения граждан	Изменение системы социального обеспечения Усиление трудовых прав работников применительно к выполнению удаленной работы и соблюдению баланса между рабочим временем и временем отдыха	Государственная служба/Позитивное государственное управление
Право на социальное обеспечение	Дискриминация Риск предвзятости алгоритмов Риск уязвимости к ошибкам Информационная асимметрия Риск стандартизации поведения граждан Цифровое неравенство Недостаточная понятность функционирования ИИ	Объяснимый ИИ; Право на объяснение; Информационное самоопределение; Изменение системы социального обеспечения; Человек в контуре управления (humans-in-the-loop) – участие человека в цикле обработки данных; Проектируемая оспоримость	Оказание государственных услуг Государственная поддержка уязвимых категорий граждан

Продолжение таблицы 2

Право человека	Риски	Способы преодоления рисков	Применение в целом и в сфере государственного управления
Право на охрану здоровья и медицинскую помощь	Дискриминация Риск предвзятости алгоритмов Риск уязвимости к ошибкам Информационная асимметрия Риск стандартизации поведения граждан	Право на объяснение; Информационное самоопределение; Проектируемая оспоримость	Оказание государственных услуг Государственная поддержка уязвимых категорий граждан
Право на благоприятную окружающую среду	Цифровое неравенство Дискриминация Риск предвзятости алгоритмов Риск уязвимости к ошибкам Информационная асимметрия	Право на доступ к Интернету; Информационное самоопределение; Подотчетность при использовании алгоритмов в процессе принятия решений в государственном секторе	Оказание государственных услуг
Право на образование	Цифровое неравенство Дискриминация Риск предвзятости алгоритмов Риск уязвимости к ошибкам Информационная асимметрия	Проведение переподготовки кадров; Переориентация и адаптация системы образования; Развитие цифровых знаний, умений и навыков	Оказание государственных услуг Государственная поддержка уязвимых категорий граждан
Свобода литературного, художественного, научного, технического и других видов творчества и преподавания	Риск предвзятости алгоритмов Риск уязвимости к ошибкам Информационная асимметрия	Право на доступ к Интернету; Информационное самоопределение; Право на забвение	Оказание государственных услуг
Право на участие в культурной жизни и пользование учреждениями культуры, на доступ к культурным ценностям	Риск предвзятости алгоритмов Риск уязвимости к ошибкам Информационная асимметрия	Право на доступ к Интернету	Оказание государственных услуг
Право на судебную защиту своих прав и свобод	Риск стандартизации поведения граждан Недостаточная понятность функционирования ИИ Дискриминация Риск предвзятости алгоритмов Риск уязвимости к ошибкам Информационная асимметрия	Право на объяснение; Информационное самоопределение Подотчетность при использовании алгоритмов в процессе принятия решений в государственном секторе; Надзор со стороны контролирующих органов и юридическая ответственность; Оценка алгоритмического воздействия; Объяснимый ИИ	Отправление правосудия

Проведенная выше систематизация может быть «специализирована» применительно к государственному управлению. Для этого рассмотрим функции государственного управления – нормотворчество, контрольно-разрешительная деятельность, оказание государственных услуг и отправление правосудия. При их осуществлении безусловно оказывается влияние на права человека, а использование ИИ приводит к определенным рискам.

Таблица 3.

Систематизация рисков нарушений наиболее значимых прав человека при использовании ИИ в реализации функций государственного управления

Функция государственного управления	Применимость ИИ	Влияние на права человека	Риски
Нормотворчество	Аналитика (обезличенные данные); Машинное обучение: - юридическая экспертиза нормативно-правовых актов; - машиночитаемое право, в том числе с помощью технологии обработки естественного языка	Право на неприкосновенность частной жизни, личную и семейную тайну, защиту своей чести и доброго имени, включая право на защиту персональных данных; Право на информацию; Право на участие в управлении делами государства	Риск стандартизации поведения граждан; Риск уязвимости к ошибкам; Риск предвзятости алгоритмов
Контрольно-разрешительная деятельность	Машинное обучение: - принятие типовых, простых административных решений Наблюдение	Право на неприкосновенность частной жизни, личную и семейную тайну, защиту своей чести и доброго имени, включая право на защиту персональных данных; Право на достоинство личности; Право на свободное использование своих способностей и имущества для предпринимательской и иной не запрещенной законом экономической деятельности; Право на информацию; Право частной собственности	Риск стандартизации поведения граждан; Информационная асимметрия; Риск уязвимости к ошибкам; Риск предвзятости алгоритмов Риск охлаждающего эффекта Чрезмерное вмешательство в частную жизнь

Продолжение таблицы 3

Функция государственного управления	Применимость ИИ	Влияние на права человека	Риски
Оказание государственных услуг	<p>Машинное обучение/ Программирование эмоций:</p> <ul style="list-style-type: none"> - принятие типовых, простых административных решений; 	<p>Право на неприкосновенность частной жизни, личную и семейную тайну, защиту своей чести и доброго имени, включая право на защиту персональных данных;</p> <p>Право на достоинство личности;</p> <p>Право на социальное обеспечение;</p> <p>Право на свободное использование своих способностей и имущества для предпринимательской и иной не запрещенной законом экономической деятельности;</p> <p>Право на информацию;</p> <p>Право частной собственности</p>	<p>Дискриминация;</p> <p>Риск стандартизации поведения граждан;</p> <p>Информационная асимметрия;</p> <p>Риск уязвимости к ошибкам;</p> <p>Риск предвзятости алгоритмов;</p> <p>Цифровое неравенство;</p> <p>Чрезмерное вмешательство в частную жизнь</p>
Правосудие	<p>Машинное обучение/ Автоматизированное принятие решений с обязательным контролем со стороны человека:</p> <ul style="list-style-type: none"> - судебная аналитика, в том числе в целях оказания содействия при формировании единой правовой позиции высших судов; - в качестве «помощника» при принятии решений по конкретным делам (простые, типовые); - алгоритмическая оценка риска совершения правонарушений в будущем (предиктивная аналитика). <p>Электронное правосудие, в том числе оборот данных между государственными органами</p>	<p>Равенство</p> <p>Право на неприкосновенность частной жизни, личную и семейную тайну, защиту своей чести и доброго имени, включая право на защиту персональных данных;</p> <p>Право на достоинство личности;</p> <p>Право на судебную защиту своих прав и свобод;</p> <p>Право на информацию;</p> <p>Право на свободу и личную неприкосновенность</p>	<p>Дискриминация;</p> <p>Риск предвзятости алгоритмов;</p> <p>Чрезмерное вмешательство в частную жизнь;</p> <p>Недостаточная понятность функционирования ИИ;</p> <p>Информационная асимметрия;</p> <p>Риск уязвимости к ошибкам;</p> <p>Цифровое неравенство</p>

Итоговый перечень представим следующим образом. На основе анализа доктрины и основополагающих документов стран, продвинутых в вопросах использования ИИ, систематизированы основные угрозы от использования ИИ в сфере государственного управления, которые «нацелены» на следующие основные права:

1. право на достоинство личности;
2. право на свободу и личную неприкосновенность
3. право на неприкосновенность частной жизни, личную и семейную тайну, защиту своей чести и доброго имени;
4. право на защиту персональных данных;
5. право на информацию;
6. право на участие в управлении делами государства
7. право частной собственности;
8. право на свободное использование своих способностей и имущества для предпринимательской и иной не запрещенной законом экономической деятельности;
9. право на социальное обеспечение;
10. право на судебную защиту своих прав и свобод.

Для наглядности сфера государственного управления представлена через основные функции – нормотворчество, контрольно-разрешительная деятельность, оказание государственных услуг и отправление правосудия. Установлено, что использование ИИ в сфере государственного управления сопряжено со следующими рисками для основных прав человека:

1. риск стандартизации поведения граждан;
2. недостаточная понятность функционирования ИИ;
3. информационная асимметрия;
4. риск предвзятости алгоритмов;
5. риск уязвимости к ошибкам;
6. риск охлаждающего эффекта;
7. чрезмерное вмешательство в частную жизнь;
8. цифровое неравенство;
9. дискриминация.

4. Правовые способы преодоления рисков нарушений прав человека при использовании ИИ в государственном управлении

Риски нарушений прав человека всегда будут оставаться одной из наиболее предполагаемых девиаций в процессе правоприменения в области государственного управления. Человек как атомистическая единица со своими интересами и государство как организованное сообщество, в котором публичный интерес формируется далеко не из простой суммы индивидуальных интересов, – вечный источник конфликта. Тем не менее, официализация и признание в государстве прав человека – серьезная опора правовой конструкции, в результате чего любое выстраивание правового регулирования (и особенно с участием новых, плохо изученных компонентов, таких как искусственный интеллект) должно принимать в расчет права человека и предусматривать механизмы их учета, нейтрализации/минимизации рисков, возмещения ущерба. Все это – элементы будущего правового регулирования использования ИИ в государственном управлении.

С содержательной точки зрения необходимо обеспечить реализацию таких концептов, которые позволяют максимально учесть возможные риски нарушений прав человека. Разумеется, основу такого подхода составляет открытость, транспарентность самого процесса государственного управления, когда и государству в лице его служащих и гражданам становятся понятны «правила игры», возможные отклонения и «плата» (компенсации) за них. Отражением транспарентности применительно к ИИ стал концепт объяснимого (понятного) ИИ. В целом правовые способы предотвращения и преодоления рисков правам человека должны выстраиваться с учетом четырех «магистральных» направлений: 1) объяснимый (понятный) ИИ, 2) сочетание публично-правовых и частно-правовых мер, 3) право на информационное самоопределение, 4) правовая детализация действий ИИ.

Сами правовые способы предотвращения и преодоления названных рисков можно представить системно.

Во-первых, в рамках разработки принципов использования ИИ в государственном управлении необходимо последовательно вводить и регламентировать новые права человека, которые позволят обеспечить встречную и превентивную защиту прав граждан при использовании в их отношении ИИ. «Человечному» использованию ИИ будут отвечать такие права, как право знать об использовании ИИ (что является «ответвлением» права на информацию), право на защиту от дискриминации со стороны ИИ (более активная форма права на недискриминацию), право на отказ от использования ИИ и сохранение «старых» отношений с государ-

ственным служащим – человеком, право на информационное самоопределение (более ответственный способ управления персональными данными). Перечисленные права составляют основу правовой конструкции использования ИИ в государственном управлении.

Во-вторых, необходимы действия по традиционному реформированию (модернизации) законодательства (внесение изменений в действующие нормативные правовые акты, отмена норм, «способствующих» нарушениям прав человека). Так, насущной является проблема урегулирования правового режима персональных данных, полученных в составе больших данных, и, следовательно, не подпадающих под сферу действия законодательства о защите персональных данных.

В-третьих, в действующее законодательство нужно вводить новые правовые конструкции на основе новых (или не новых, но еще не реализованных) концептов. Описанное в исследовании возрастание роли самого субъекта данных в управлении ими отражает право на информационное самоопределение, отсутствующее в российском законодательстве. Заслуживает внимания концепт минимизации вреда правам человека и его возмещения в гражданско-правовом порядке, концепция объяснимого ИИ. При этом желательно, чтобы сначала новые концепты появились в программно-ориентирующих актах, предусматривающих их гармоничное введение в законодательство.

В-четвертых, необходимо сочетание государственно-правовой регламентации и саморегулирования (особенно это актуально в плане этики ИИ). Для последовательной реализации человекоцентричности необходимо ввести обязательную оценку алгоритмов ИИ, применяемых в государственном управлении, на предмет потенциальных нарушений прав человека, а в процессе эксплуатации – соответствующий мониторинг. Данная оценка должна учитывать, в том числе, выделенные нами риски - риск стандартизации поведения граждан; недостаточная понятность функционирования ИИ; информационная асимметрия; риск предвзятости алгоритмов; риск уязвимости к ошибкам; риск охлаждающего эффекта; чрезмерное вмешательство в частную жизнь; цифровое неравенство; дискриминация. Для разработки как самих алгоритмов ИИ, применяемых в государственном управлении, так и правил их оценки считаем возможным использовать предложенные нами таблицы систематизации рисков.

В-пятых, организационно-правовые и процессуальные меры. Это - создание независимого органа по защите данных, расширение компетенции судов в связи со спорами, связанными с нарушением прав человека в области использования цифровых технологий. Сама система защиты прав граждан при этом должна выстраиваться с учетом потенциальных рисков: например, наличие охлаждающего эффекта должно компенсироваться более прогрессивными

процессуальными нормами о возмещении причиненного вреда. Отдельную задачу представляет «совмещение» технологических принципов функционирования ИИ с правовыми (а также этическими, социальными нормами и пр.).

Заключение

Проведенное исследование позволяет констатировать повышенный интерес к использованию ИИ в интересах, в том числе, государственного управления в разных странах мира и, одновременно, практически полное отсутствие его нормативного регулирования.

Российский опыт правового регулирования ИИ представляется наименее сформированным, характеризуемым недостаточным вниманием к концепции прав человека и увлечением технической составляющей освоения ИИ.

Основные риски от использования ИИ «нацелены» на такие основные права, как право на уважение частной жизни, защиту персональных данных, право на информацию, право на собственность, в частности, авторское право, свобода предпринимательства, права на эффективные средства правовой защиты и беспристрастное судебное разбирательство, а также право на недискриминацию.

В ходе настоящего исследования выявлен ряд рисков правам человека от использования ИИ.

Во-первых, риск стандартизации поведения, в результате реализации которой индивидуальное поведение большого количества людей будет выровнено и более единообразно. В итоге реализация индивидуальности может потребовать больших усилий и больших затрат и даже привести к социальному разделению.

Во-вторых, недостаточная понятность функционирования ИИ. Автоматизированные системы могут определять характеристики, склонности, цели и намерения людей в ранее неизвестной глубине и детализации и, таким образом, вычислять прогнозы относительно их будущего поведения.

В-третьих, информационная асимметрия. Использование автоматизированных систем может привести к информационной асимметрии, в которой преимущество одного лица возникает именно благодаря информации о и/или против другого лица. Использование автоматизированных систем дает возможность оценить, какие индивидуальные потребности и цели есть или могут быть обеспечены, независимо от того, осознает ли сам гражданин свои собственные потребности и цели.

В качестве направлений для разработки правовых способов предотвращения и преодоления рисков правам человека предложены: 1) объяснимый (понятный) ИИ, 2) сочетание публично-правовых и частно-правовых мер, 3) право на информационное самоопределение, 4) правовая детализация действий ИИ.

Сами правовые способы предотвращения и преодоления названных рисков представлены системно. В частности, новые права человека способны обеспечить встречную и

превентивную защиту прав граждан при использовании в их отношении ИИ (право знать об использовании ИИ (что является «ответвлением» права на информацию), право на защиту от дискриминации со стороны ИИ (более активная форма права на недискриминацию), право на отказ от использования ИИ и сохранение «старых» отношений с государственным служащим – человеком, право на информационное самоопределение (более ответственный способ управления персональными данными)). Необходимо решать проблему урегулирования правового режима персональных данных, полученных в составе больших данных, и, следовательно, не подпадающих под сферу действия законодательства о защите персональных данных. В действующее законодательство нужно вводить новые правовые конструкции на основе новых (или не новых, но еще не реализованных) концептов - информационное самоопределение, минимизация вреда правам человека и его возмещения в гражданско-правовом порядке, концепция объяснимого ИИ.

Для последовательной реализации человекоцентричности необходимо ввести обязательную оценку алгоритмов ИИ, применяемых в государственном управлении, на предмет потенциальных нарушений прав человека, а в процессе эксплуатации – соответствующий мониторинг. Данная оценка должна учитывать, в том числе, выделенные в настоящем исследовании риски - риск стандартизации поведения граждан; недостаточная понятность функционирования ИИ; информационная асимметрия; риск предвзятости алгоритмов; риск уязвимости к ошибкам; риск охлаждающего эффекта; чрезмерное вмешательство в частную жизнь; цифровое неравенство; дискриминация.

Не обойтись без организационно-правовых и процессуальных мер (создание независимого органа по защите данных, расширение компетенции судов в связи со спорами, связанными с нарушением прав человека в области использования цифровых технологий).

Новизна настоящего исследования определяется неурегулированностью процесса использования искусственного интеллекта в государственном управлении, что влечет риски возможности нарушений прав человека. Его результаты позволяют системно представить риски в области прав человека в условиях цифрового государственного управления, что помогает их учесть и/или нейтрализовать при разработке нормативных правовых актов.

Благодарности

Материал подготовлен в рамках выполнения научно-исследовательской работы государственного задания РАНХиГС на 2022 год по научному направлению «Государственное управление и государственная служба. Реформа государственного управления на основе развития проектного и процессного подходов».

СПИСОК ИСТОЧНИКОВ

1. The National Artificial Intelligence Initiative Act of 2020 (NAIIA). – URL: <https://www.congress.gov/116/crpt/hrpt617/CRPT-116hrpt617.pdf#page=1210> (дата обращения 10.05.2022).
2. Consolidated Appropriations Act, 2021 <https://www.congress.gov/116/bills/hr133/BILLS-116hr133enr.pdf#page=1105> (дата обращения 10.05.2022).
3. Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government. – URL: <https://www.federalregister.gov/documents/2020/12/08/2020-27065/promoting-the-use-of-trustworthy-artificial-intelligence-in-the-federal-government> (дата обращения 10.05.2022).
4. Legislation and Executive Orders. – URL: <https://www.ai.gov/legislation-and-executive-orders/> (дата обращения 10.05.2022).
5. White House Proposes Tech 'Bill of Rights' to Limit AI Harms. – URL: <https://www.usnews.com/news/business/articles/2021-10-08/white-house-proposes-tech-bill-of-rights-to-limit-ai-harms> (дата обращения 05.05.2022).
6. Report on artificial intelligence in a digital age (2020/2266(INI)). Special Committee on Artificial Intelligence in a Digital Age. – URL: <https://www.europarl.europa.eu/committees/en/aida/documents/latest-documents> (дата обращения 11.05.2022).
7. China Innovation Project. A Guide to Understanding China's Next Wave of Innovation. – URL: <https://projects.iq.harvard.edu/innovation/made-china-2025-explained#:~:text=MIC%202025%20is%20an%20initiative,compete%20both%20domestically%20and%20globally> (дата обращения 05.05.2022).
8. Ковачич Л. Китайский опыт развития отрасли искусственного интеллекта: стратегический. – URL: <https://carnegieendowment.org/2020/07/07/ru-pub-82172> (дата обращения 05.05.2022).
9. Coordinated Plan on Artificial Intelligence. – URL: <https://digital-strategy.ec.europa.eu/en/policies/plan-ai#:~:text=The%20key%20aims%20of%20the,AI%20policy%20to%20avoid%20fragmentation.&text=The%20Coordinated%20Plan%20on%20Artificial%20Intelligence%202021%20Review%20is%20the,global%20leadership%20in%20trustworthy%20AI.> (дата обращения 11.05.2022).
10. Proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts (COM(2021)0206). – URL: <https://eur-lex.europa.eu/legal-content/EN/TXT/DOC/?uri=CELEX:52021PC0206&from=EN> (дата обращения 11.05.2022).
11. Организационно-правовые проблемы АСУ. Отв. ред. И.Л. Бачило. М., 1979.

12. Дорожная карта развития «Сквозной» цифровой технологии «Нейротехнологии и искусственный интеллект». – URL: <https://digital.gov.ru/uploaded/files/07102019ii.pdf> (дата обращения 19.08.2022).
13. Ernst C. Artificial Intelligence and Autonomy: Self-Determination in the Age of Automated Systems // Th. Wischmeyer, T. Rademacher (Eds.). *Regulating Artificial Intelligence*. Springer, Cham, 2020. 388 p.
14. Rademacher T. Artificial Intelligence and Law Enforcement // Th. Wischmeyer, T. Rademacher (Eds.). *Regulating Artificial Intelligence*. Springer, Cham, 2020. P. 226.
15. Пибяев И.А., Симонова С.В. Алгоритмы в механизме реализации конституционных прав и свобод: вызовы цифровой эпохи // *Сравнительное конституционное обозрение*. 2020. № 6. С. 31 - 50.
16. Черешнева И.А. Алгоритмы и право: анализ, вдохновленный судебным прецедентом // *Информационное общество*. 2021. № 6. С. 84-90.
17. Rowena R. (2020). Legal and human rights issues of AI: Gaps, challenges and vulnerabilities. – URL: <https://www.sciencedirect.com/science/article/pii/S2666659620300056> (дата обращения 27.04.2022).

**В СЕРИИ ПРЕПРИНТОВ
РАНХиГС РАССМАТРИВАЮТСЯ
ТЕОРЕТИЧЕСКИЕ
И ПРАКТИЧЕСКИЕ ПОДХОДЫ
К СОЗДАНИЮ, АКТИВНОМУ
ИСПОЛЬЗОВАНИЮ
ВОЗМОЖНОСТЕЙ
ИННОВАЦИЙ В РАЗЛИЧНЫХ
СФЕРАХ ЭКОНОМИКИ
КАК КЛЮЧЕВОГО УСЛОВИЯ
ЭФФЕКТИВНОГО УПРАВЛЕНИЯ**



РАНХиГС
РОССИЙСКАЯ АКАДЕМИЯ НАРОДНОГО ХОЗЯЙСТВА
И ГОСУДАРСТВЕННОЙ СЛУЖБЫ
ПРИ ПРЕЗИДЕНТЕ РОССИЙСКОЙ ФЕДЕРАЦИИ