

Политика и Общество

Правильная ссылка на статью:

Косоруков А.А. Большие данные и искусственный интеллект в государственном управлении: возможности и риски // Политика и Общество. 2025. № 4. DOI: 10.7256/2454-0684.2025.4.73363 EDN: XIPDGV URL: https://nbpublish.com/library_read_article.php?id=73363

Большие данные и искусственный интеллект в государственном управлении: возможности и риски

Косоруков Артем Андреевич

ORCID: 0000-0002-0275-4899

кандидат политических наук

доцент, кафедра политического анализа; Московский государственный университет имени М.В.Ломоносова

119991, Россия, г. Москва, ул. Ломоносовский Проспект, 27/4, оф. Г645

✉ kosorukov@spa.msu.ru



[Статья из рубрики "АКТУАЛЬНЫЙ ВОПРОС"](#)

DOI:

10.7256/2454-0684.2025.4.73363

EDN:

XIPDGV

Дата направления статьи в редакцию:

15-02-2025

Аннотация: Статья посвящена анализу возможностей и рисков использования больших данных и искусственного интеллекта (ИИ) в государственном управлении. Автор рассматривает концепцию больших данных, их ключевые характеристики (объем, скорость, разнообразие, достоверность) и способы их обработки с помощью ИИ. Основное внимание уделяется применению ИИ-систем в таких сферах, как прогнозирование социально-экономического развития, здравоохранение, образование, борьба с преступностью и оказание государственных услуг. Рассматриваются примеры успешного внедрения ИИ в разных странах (Сингапур, США, Китай, Россия, Великобритания, Южная Корея). Также анализируются риски, связанные с цифровой предвзятостью, киберпреступностью, вмешательством в частную жизнь, социальной несправедливостью и дезинтеграцией социальных связей. Автор подчеркивает важность регулирования использования ИИ в государственном и корпоративном секторах, включая юридическую ответственность разработчиков, корпораций и госорганов за

предвзятые алгоритмические решения, а также внедрение этических стандартов и механизмов обратной связи от пользователей. Методология исследования включает в себя экспертный опрос с использованием онлайн-таблицы на платформе Google, позволившей проанкетировать 24 представителя академического сообщества России в течение октября–декабря 2024 г. Респондентами выступили ведущие российские эксперты из 16 федеральных вузов России, представляющих все 8 федеральных округов (по 3 эксперта от каждого округа). Научная новизна исследования заключается в систематизации рисков, связанных с внедрением больших данных и ИИ в государственное управление, на основе экспертного анкетирования ведущих российских специалистов из 8 федеральных округов. Впервые предложена ранжированная классификация рисков, включая цифровую предвзятость, нетранспарентность решений ИИ, угрозу персональным данным и возможное усиление социальной несправедливости. Выводы статьи подчеркивают необходимость регулирования ИИ-алгоритмов, внедрения этических стандартов и разработки механизмов защиты граждан от дискриминационных решений. Автор подчеркивает важность тестирования алгоритмов на предвзятость, корпоративной ответственности разработчиков и обратной связи от пользователей. Исследование вносит вклад в развитие концепции цифрового государственного управления, демонстрируя баланс между технологическими возможностями и потенциальными рисками от внедрения ИИ и больших данных.

Ключевые слова:

Большие данные, Искусственный интеллект, Государственное управление, Машинное обучение, Прогнозирование, Киберпреступление, Цифровая предвзятость, Персональные данные, Автоматизация, Этические стандарты

Введение

Большие данные представляют собой концепцию, охватывающую объемы, скорость и разнообразие данных, которые не могут быть эффективно обработаны традиционными архитектурами обработки и стандартными техническими средствами. Концепция больших данных начала развиваться в 1990-х годах, когда начали появляться первые системы для обработки больших объемов информации. Однако термин «большие данные» (Big Data) впервые был использован в научном контексте в начале 2000-х годов. Одним из первых, кто применил этот термин, был исследователь Дуглас Лэмперт, который описал большие данные как данные, которые характеризуются тремя «V»: объемом (Volume), скоростью (Velocity) и разнообразием (Variety). Растущий объем (Volume) данных, генерируемых в различных сферах, таких как социальные сети, IoT (интернет вещей), финансовые транзакции и другие, составляет терабайты и петабайты, требует разработки новых архитектур хранения и обработки данных. Данные поступают с высокой скоростью (Velocity), что требует оперативной обработки в реальном времени, особенно для таких приложений, как мониторинг финансовых рынков или управление производственными процессами. Помимо этого, большие данные описываются термином «разнообразие» (Variety), так как они могут быть структурированными, полуструктурированными и неструктурированными, требуют применения различных методов анализа, включая текстовый анализ, обработку изображений и видео, а также термином «достоверность» (Veracity), ведь качество и надежность данных могут варьироваться, что ставит задачу обеспечения достоверности информации и ее источников. При этом извлечение

полезной информации из больших объемов данных требует применения сложных аналитических методов и алгоритмов, таких как машинное обучение и искусственный интеллект.

Возможности использования больших данных и искусственного интеллекта в государственном управлении

В 2005 году термин «большие данные» стал набирать популярность благодаря выступлению таких дата-специалистов, как Бен Шнайдерман и других, которые начали активно обсуждать возможности и вызовы, связанные с обработкой больших данных. С тех пор термин «большие данные» стал широко использоваться в научных исследованиях, бизнесе и различных отраслях, что привело к значительному развитию технологий и методов анализа данных [\[1\]](#).

Исследуя возможности больших данных в современном государственном управлении, следует учитывать, что они могут поступать из различных источников и иметь различный формат, в частности:

1. Сенсоры и устройства Интернета вещей (IoT), установленные в городах, на транспорте и в других местах (например, умные счетчики, системы мониторинга окружающей среды, датчики трафика), собирают данные о различных параметрах (например, температуре, влажности, движении и т.д.);
2. Веб-скрейпинг- технология, позволяющая автоматически извлекать данные с веб-сайтов, чаще всего используется для сбора информации из открытых источников, таких как социальные медиа, новостные сайты и т.д., может применяться в различных областях, включая маркетинг, исследование рынка, анализ конкурентов, сбор данных для машинного обучения и многое другое;
3. Социальные сети выступают одним из важнейших источников больших данных, а данные о пользователях и их взаимодействиях могут быть собраны через API социальных сетей или с помощью анализа контента. Социальные сети предоставляют API, которые позволяют разработчикам получать доступ к различным данным о пользователях и их взаимодействиях: обработка естественного языка (NLP) поддерживает анализ текстов постов и комментариев, лежит в основе сентимент-анализа - оценки эмоциональной окраски текстов, помогая понять, как пользователи реагируют на определенные темы или события; компьютерное зрение с использованием таких библиотек, как OpenCV или TensorFlow, применяется для анализа изображений, распознавания объектов, лиц или сцен; анализ взаимодействий между пользователями применяется для построения социальных графов, позволяющих выявлять влиятельных пользователей и сообщества, проводить их кластеризацию и сегментацию [\[2\]](#).
4. Онлайн-опросы и анкеты как инструмент сбора больших данных о мнениях и предпочтениях граждан (например, Google Forms - бесплатное приложение от Google, которое позволяет создавать и масштабировать онлайн-опросы и анкеты, используя широкую номенклатуру вопросов, логик ветвления, визуализации данных и интеграции с другими сервисами, Survey Monkey - одна из наиболее известных платформ для создания и проведения онлайн-опросов, предлагающая широкий функционал, включая продвинутые аналитические инструменты, сегментацию аудитории и интеграцию с

внешними системами).

5. Ведомственная статистика и базы данных, платформы государственных услуг, собирающие большие данные, касающиеся как отдельных граждан, так и показателей экономики, демографии, здравоохранения, транспорта, безопасности и др.

В связи с объемом и сложностью больших данных, постоянным процессом их производства, их эффективный анализ в современном государственном управлении практически невозможен без применения искусственного интеллекта. Информационные системы на базе искусственного интеллекта (далее ИИ-системы) могут обрабатывать и анализировать огромные объемы данных из различных источников, выявляя скрытые закономерности и тренды. Алгоритмы машинного обучения способны строить более точные модели и прогнозы, чем традиционные статистические методы, особенно в отношении сложных, нелинейных процессов [3]. ИИ-системы автоматизируют сбор и первичную обработку данных, высвобождая время госслужащих для аналитической работы. Прогнозные модели ИИ помогают лицам, принимающим решения, оценивать последствия различных сценариев и выбирать оптимальные варианты. Системы ИИ способны быстрее выявлять признаки надвигающихся кризисов, стихийных бедствий, эпидемий и других проблем. ИИ-системы могут непрерывно обучаться и совершенствовать свои модели по мере поступления новых данных. Применения ИИ и больших данных в государственном управлении лежит в основе моделирования и прогнозирования социально-экономических процессов, финансового рынка и рынка труда, демографических показателей, чрезвычайных ситуаций, выявления рисков и угроз, обнаружения кризисных явлений на их латентной стадии [4]. В частности, система прогнозирования социально-экономического развития в Сингапуре (SPIS) использует машинное обучение для анализа больших данных из различных источников и построения прогнозных моделей по ключевым экономическим и социальным показателям государства. Система прогнозирования преступности в США (PredPol) использует алгоритмы машинного обучения для анализа данных о преступлениях и построения прогнозных моделей, что позволяет полиции предотвращать правонарушения. Система раннего предупреждения о стихийных бедствиях в КНР (EWDS) использует данные с датчиков, спутников, социальных сетей и другие источники для мониторинга и прогнозирования природных катастроф.

На основе анализа больших данных системы искусственного интеллекта могут предоставлять государственным служащим рекомендации для принятия решений, касающихся бюджета, социальных программ и других управленческих задач, что помогает оптимизировать распределение ресурсов. ИИ-системы могут отслеживать эффективность реализуемых социальных программ и инициатив, позволяя вносить коррективы в режиме реального времени. ИИ-систему могут анализировать заявки на социальные выплаты и выявлять аномалии в паттернах поведения, позволяя обнаруживать потенциальные случаи мошенничества. Системы могут автоматически сигнализировать о подозрительных заявках, поддерживая государственные органы в вопросах проведения более глубоких проверок. Использование ИИ для обнаружения мошенничества позволяет сократить потери бюджета и направить средства на поддержку действительно нуждающихся граждан. ИИ может анализировать отзывы граждан, их клиентский опыт взаимодействия с порталами госуслуг, выявлять общие проблемы или запросы, помогая государственным ведомствам улучшать качество предоставляемых услуг и уровень удовлетворенности со стороны населения.

Среди разнообразных направлений применения больших данных и систем

искусственного интеллекта в современном государственном управлении следует выделить процесс оказания государственных услуг, так как он охватывает практически все население и затрагивает сферу наиболее чувствительных личных данных граждан. Так, в Российской Федерации активно внедряются чат-боты, в основе которых лежат обученные на больших данных нейросети, позволяющие проводить общие консультации и информирование граждан, повышая доступность и скорость обслуживания населения. Например, виртуальные помощники столичных МФЦ, виртуальные ассистенты в Росреестре, ФНС, Пенсионном фонде обеспечивают доступность своих услуг 24/7, предоставляют персонализированные ответы, снижают нагрузку на сотрудников. Системы ИИ используются для автоматизации рутинных административных процессов, лежащих в основе оказания госуслуг, повышая эффективность и сокращая трудозатраты [\[5\]](#). Так, в Республике Корея в 2012 году внедрена система автоматического начисления пособий, основанная на анализе данных гражданина. Активно разрабатываются интеллектуальные системы, которые помогают государственным служащим принимать более обоснованные решения, анализируя большие объемы данных. Например, в Великобритании уже более 10 лет используется система поддержки принятия решений по предоставлению социальных пособий, учитывающая множество факторов заявителя.

В области здравоохранения ИИ-системы анализируют рентгеновские снимки другие медицинские изображения для выявления признаков различных заболеваний, таких как рак, заболевания сердца, неврологические расстройства и многое другое, помогая врачам поставить точный диагноз на ранней стадии. Алгоритмы машинного обучения таких систем как Botkin.AI, RadIO, IBMWatsonHealth и др., обученные на больших наборах данных, способны выявлять мельчайшие детали и паттерны, которые могут быть незаметны для человеческого глаза, позволяя повысить точность диагностики по сравнению с традиционными методами. Модели ИИ могут анализировать большие объемы данных о пациентах, включая медицинские записи, результаты анализов, образ жизни и другую информацию, чтобы выявлять закономерности и факторы риска развития определенных заболеваний. На основе этого анализа ИИ-системы способны предсказывать вероятность развития заболеваний у конкретных пациентов, что помогает врачам принимать профилактические меры и разрабатывать индивидуальные планы ухода. Возможность предсказывать развитие заболеваний позволяет врачам вмешиваться на ранних стадиях, что значительно улучшает прогноз для пациентов. ИИ-системы могут обрабатывать и анализировать геномные данные пациентов, чтобы выявлять генетические маркеры, связанные с определенными заболеваниями или реакцией на лекарственные препараты. На основе этого анализа ИИ может помогать врачам разрабатывать индивидуальные планы лечения, учитывающие генетические особенности, состояние здоровья и другие факторы конкретного пациента [\[6\]](#).

В сфере образования ИИ-системы могут анализировать уровень знаний, предпочтения и стиль обучения каждого студента. На основе больших данных они адаптируют учебные материалы, предлагая подходящие задания, тесты и ресурсы. Данные системы, в частности, Smart Sparrow, Gradescope, LearningAnalytics и др. могут рекомендовать студентам дополнительные ресурсы, такие как видео-лекции, статьи или интерактивные упражнения, которые помогут углубить понимание тем, с которыми чаще всего возникают трудности. Если студент справляется с материалом, ИИ предлагает более сложные задачи, в то время как для тех, кто испытывает трудности, он может предложить дополнительные пояснения или более простые задания. ИИ может анализировать данные об успеваемости студентов, включая результаты тестов, участие в занятиях и выполнение заданий. На основе анализа успеваемости ИИ может предлагать соответствующие ресурсы, такие как дополнительные занятия, учебные материалы или

консультации с преподавателями. ИИ может отслеживать прогресс студентов в реальном времени, предоставляя преподавателям информацию о том, как каждый студент справляется с материалом. В конечном итоге, ИИ может использоваться для автоматической оценки письменных работ и тестов, значительно сокращая время, затрачиваемое учителями на проверку, и снижая вероятность субъективности в оценивании. Перспективным представляется внедрение в учебный процесс виртуальных ИИ-помощников, которые могут отвечать на вопросы студентов в режиме реального времени, предоставляя информацию о курсах, расписаниях, учебных материалах и других ресурсах на основе их интересов и потребностей, что способствует более эффективному обучению [\[7\]](#).

Методология исследования

Несмотря на обозначенные преимущества, важно раскрыть ряд потенциальных рисков при использовании больших данных и искусственного интеллекта в государственном управлении. Однако исследование рисков внедрения искусственного интеллекта в государственное управление во многом лежит в плоскости экспертных опросов, в связи с чем автором было проведено анкетирование экспертного академического сообщества России в формате заполнения онлайн-таблицы с применением платформы Google (октябрь–декабрь 2024 г.). 24 респондентов, которые являются ведущими российскими экспертами из федеральных вузов нашей страны, представляющих все 8 федеральных округов (3 эксперта из каждого округа), ознакомили с целями данного исследования и попросили выделить наиболее вероятные риски от внедрения искусственного интеллекта в сферу государственного управления. Ранжированный список рисков, указанных экспертами, представлен в Таблице 1.

Таблица 1. Риски от внедрения искусственного интеллекта в государственное управление

№	Риски от внедрения искусственного интеллекта	Количество экспертов, указавших данный риск под № 1 или 2
1	Киберпреступления в отношении данных и систем	13
2	Нетранспарентность и непредсказуемость решений	11
3	Вмешательство в частную жизнь граждан	9
4	Дискриминация человека по различным основаниям	7
5	Усиления социальной несправедливости	6
6	Изоляции и дезинтеграции социальных связей	2

Риски использования искусственного интеллекта и больших данных в государственном управлении

Рассмотрим каждый из рисков, указанных в таблице 1, подробнее.

Использование искусственного интеллекта в государственном управлении несет серьезные риски, связанные с киберпреступлениями в отношении ведомственных и персональных данных, информационных систем и хранилищ. Как и любое программное обеспечение, ИИ-системы могут содержать уязвимости, которые могут быть использованы злоумышленниками для взлома, кражи данных или вывода систем из строя (в частности, при помощи атак с применением машинного обучения, Deepfake-мошенничества, ИИ-ботнетов, адаптивных DDoS-атак и др.). Сами ИИ-системы могут быть использованы для автоматизации и масштабирования кибератак, таких как массовая рассылка вредоносного программного обеспечения, фишинговые атаки или взлом систем. Злоумышленники могут пытаться внедрять ложные или искаженные данные в ИИ-системы, чтобы влиять на их решения в своих интересах. Преступники могут применять ИИ-технологии для создания более изощренных и трудно обнаруживаемых кибератак, например, для генерации вредоносного программного обеспечения, адаптированного под конкретные системы. Централизованные ИИ-системы в государственном управлении могут стать привлекательными целями для кибератак, ведущих к массовому нарушению режимов конфиденциальности в отношении персональных данных граждан [\[8\]](#).

Одними из важнейших рисков использования ИИ в государственном управлении выступают нетранспарентность и непредсказуемость решений ИИ-систем. Многие современные ИИ-системы, основанные на глубоком машинном обучении, именуются «черными ящиками», когда их внутренние алгоритмические механизмы являются чрезвычайно сложными и не до конца понятными даже разработчикам. Это создает серьезные проблемы, если такие системы начинают применяться для принятия важных государственных решений, затрагивающих права и интересы граждан. Необоснованные или дискриминационные результаты могут негативно сказываться на таких ключевых сферах как принятие государственных решений, распределение государственных ресурсов и оказание услуг, снижать доверие к правоприменительной деятельности и судопроизводству, государственному надзору и контролю. Кроме того, отсутствие возможности проследить логику ИИ-системы затрудняет или делает невозможным привлечение виновных лиц к ответственности в случае принятия неправомерных решений и последующих действий [\[9\]](#).

Серьезным риском использования искусственного интеллекта в государственном управлении является возможность вмешательства в частную жизнь граждан. Многие современные ИИ-системы, применяемые в государственном секторе, требуют обработки больших объемов персональных данных - от персональных характеристик и финансовых показателей до данных о местоположении, содержании электронных коммуникаций и поведенческих паттернов человека. Алгоритмы машинного обучения могут анализировать эту информацию для принятия решений в различных сферах: правоохранительная деятельность (профилирование, слежка, предотвращение преступлений), предоставление государственных услуг (оценка нуждаемости, назначение пособий и льгот), регулирование общественной жизни (контроль за передвижениями, социальные рейтинги). Однако тотальный мониторинг и сбор информации о гражданах может вести к грубому нарушению права на неприкосновенность частной жизни, злоупотреблениям и манипуляциям со стороны властей. Кроме того, уязвимости в системах хранения и обработки данных могут приводить к утечкам и незаконному использованию личной информации.

Если алгоритмы ИИ были обучены на больших данных, отражающих предвзятость человеческих решений, они могут воспроизводить и усиливать дискриминацию по половому, расовому, возрастному и др. признакам при найме на работу в государственные органы. ИИ-системы, основанные на неполных или искаженных данных, могут принимать решения, ущемляющие права граждан из определенных социальных групп. Например, применение ИИ в правоохранительной деятельности может приводить к избыточному и необоснованному контролю за отдельными категориями населения, выделенными при помощи интеграции больших данных, что нарушает права на неприкосновенность частной жизни [\[10\]](#). Использование ИИ в государственном управлении несет в себе риски, связанные с правами и свободами граждан, отказом от автономии личности, права на регресс и других прав. Чрезмерное применение ИИ-систем для принятия решений вместо человека может привести к ограничению права граждан на самостоятельный выбор и принятие решений, касающихся их жизненного благополучия. Если решения, принятые ИИ-системами, будут необоснованными или ошибочными, граждане могут быть лишены возможности обжаловать их в суде или добиться возмещения ущерба. Использование ИИ-систем для мониторинга и цензуры контента в интернете может создать условия для подавления свободы слова и доступа к информации.

Использование искусственного интеллекта в государственном управлении несет риск усиления социальной несправедливости. Если большие данные, используемые для обучения ИИ-систем, содержат исторические предубеждения и дискриминацию, это может привести к воспроизведению и усилению данных предвзятостей в принимаемых системами решениях. Многие ИИ-алгоритмы работают по нечетким моделям, что затрудняет понимание логики их решений, снижает подотчетность и возможность оспаривания несправедливых решений. Централизованное использование искусственного интеллекта в госуправлении может привести к концентрации власти и контроля в руках узкого круга лиц, что создает риски злоупотребления и манипулирования. Внедрение ИИ-технологий в госсекторе может усилить разрыв между теми, кто имеет доступ к цифровым ресурсам, и теми, кто его лишен, приводя к углублению социального неравенства. Полная автоматизация принятия решений государственными органами с помощью ИИ может лишить граждан возможности апелляции и индивидуального рассмотрения их заявлений и жизненных ситуаций в целом.

Использование искусственного интеллекта в государственном управлении может также привести к изоляции и дезинтеграции социальных связей граждан. Когда ИИ-системы начинают самостоятельно принимать ключевые решения в социальной сфере (например, назначать социальные пособия, зарплаты и пенсии, выделять социальное жилье), это сводит к минимуму непосредственное общение граждан с представителями органов государственного управления, что, в свою очередь, приводит к ощущению отдаленности и отчужденности граждан от государственных структур (с одной стороны, это положительно сказывается на снижении коррупции, с другой стороны, устраняется возможность живого общения, выступающая одним из способов преодоления недостатков любых административных алгоритмов и процедур). ИИ-алгоритмы, ориентированные на обезличенные статистические данные, могут недооценивать или игнорировать специфические потребности отдельных категорий граждан - пожилых людей, безработных, лиц с ОВЗ, социальных маргиналов, усугубляя их социальную исключенность. Внедрение государственных ИИ-систем, призванных контролировать и регулировать общественные отношения (социальные рейтинги, ограничения передвижений и коммуникаций), может разрушать традиционные формы взаимопомощи

локальных сообществ и неформальные социальных связей. Закрытость «черных ящиков» ИИ-систем, используемых государством, снижает возможности граждан влиять на решения, затрагивающие их жизнь.

Внедрение больших данных в государственное управление актуализирует однозначный риск, выступающий производной от ИИ-технологий и систем обеспечения цифровой безопасности, а именно – риск цифровой предвзятости. Цифровая предвзятость возникает, когда алгоритмы ИИ принимают решения, основываясь на предвзятых данных и алгоритмических моделях, что приводит к несправедливым или дискриминационным результатам, включая дискриминацию по расе, полу, возрасту, социальному статусу и другим характеристикам. Цифровая предвзятость как систематическая ошибка, возникающая в процессе создания алгоритмических моделей, приводит к дискриминации или неравному обращению с определенными группами населения [\[11\]](#). Проявляться цифровая предвзятость может в различных сферах. Алгоритмы, используемые для автоматизации процесса найма, могут дискриминировать женщин, если они обучены на статистических данных, где преобладают мужчины. Например, один из алгоритмов Amazon, разработанный для автоматизации подбора резюме, был закрыт, когда стало известно, что он дискриминирует женщин, так как обучался на данных, в которых преобладали резюме мужчин. Алгоритмы, использующиеся для таргетирования рекламы, могут игнорировать определенные возрастные группы. Например, реклама для финансовых продуктов может быть нацелена в основном на молодежь, что может исключать пожилых людей, которые также могут быть заинтересованы в таких продуктах. Некоторые кредитные скоринговые модели могут учитывать возраст как фактор, что может негативно сказываться на молодых заемщиках, у которых нет долгой кредитной истории. Использование алгоритмов для оценки кредитоспособности может привести к дискриминации людей с низким социально-экономическим статусом. Например, алгоритмы могут учитывать адрес проживания как фактор, что может негативно сказаться на людях, живущих в неблагополучных районах. В некоторых случаях алгоритмы могут ограничивать доступ к определенным услугам (например, медицинским или образовательным) для людей с низким доходом, основываясь на данных, которые не учитывают индивидуальные обстоятельства [\[12\]](#).

Когда люди осознают, что алгоритмы принимают предвзятые решения, это может привести к потере доверия к государственным учреждениям, финансовым организациям и другим системам, использующим такие технологии. Цифровая предвзятость может усугубить социальное и экономическое неравенство, создавая дополнительные барьеры для определенных групп населения [\[13\]](#). Алгоритмы, обученные на статистических данных, могут унаследовать предвзятости, существующие в этих данных, что может привести к дискриминации определенных групп, например, в области кредитования, трудоустройства или правоприменения. Несправедливые алгоритмические решения могут привести к маргинализации определенных групп, что ограничивает их доступ к ресурсам и возможностям. В свою очередь, это может создать порочный круг, где маргинализированные группы становятся еще более уязвимыми и изолированными. Когда определенные группы систематически подвергаются дискриминации, это может привести к ухудшению их экономического положения, что в свою очередь создает дополнительные барьеры для выхода из бедности и неравенства. Алгоритмы, используемые в образовательных системах, могут предвзято оценивать способности студентов, что может привести к неравному распределению образовательных ресурсов и возможностей, усугубляя образовательное неравенство и ограничивая социальную мобильность. Алгоритмические системы, используемые для отбора кандидатов на работу, могут

неосознанно дискриминировать определенные группы, что снижает их шансы на трудоустройство и карьерный рост. Углубление социального неравенства может привести к снижению доходов у маргинализированных групп, что в свою очередь снижает их покупательную способность и экономическую активность. Увеличение неравенства может привести к росту социальной напряженности и конфликтов, что негативно сказывается на экономическом развитии и стабильности общества.

Заключение

Несмотря на то, что большие данные и искусственный интеллект активно внедряются в сферу государственного управления, ключевой проблемой выступает определение юридической ответственности за предвзятые цифровые решения, принимаемые с использованием искусственного интеллекта. Вопросы ответственности затрагивают как разработчиков технологий, так и государственные органы. Разработчики ИИ должны обеспечивать, чтобы алгоритмы были протестированы на наличие предвзятости и других потенциальных проблем. Если алгоритм приводит к дискриминационным результатам из-за недостатков в его разработке, разработчики должны нести ответственность. Разработчики должны следовать этическим стандартам и лучшим практикам в области разработки ИИ. Невыполнение этих стандартов может привести к имиджевым и даже политическим последствиям, если речь идет о крупных компаниях и государственных структурах. Государственные органы, использующие ИИ для принятия решений, также несут ответственность за то, как они применяют эти технологии. Если они используют алгоритмы, которые не были должным образом проверены или которые имеют известные недостатки, это может привести к общественному резонансу, административным и кадровым решениям. Государственные органы должны проводить регулярные оценки и мониторинг используемых алгоритмов, чтобы убедиться в их справедливости и отсутствии цифровой предвзятости. Невыполнение этих обязанностей может привести к персональной ответственности их руководителей. В некоторых случаях может быть уместно установить режим совместной ответственности, где как разработчики, так и администраторы ИИ-систем несут свою долю ответственности за предвзятые решения. В договорах между разработчиками и государственными органами могут быть прописаны условия, касающиеся ответственности за цифровую предвзятость и другие нарушения, связанные с использованием ИИ [\[14\]](#).

Принятие корпоративных этических стандартов для разработчиков искусственного интеллекта, направленных против цифровой предвзятости, становится распространенной практикой в последние годы [\[15\]](#). Существует ряд примеров компаний, которые внедрили такие стандарты: 1. Google AI Principles. В 2018 году Google опубликовала свои «Принципы ИИ», которые определяют этические стандарты для разработки и использования ИИ. Принципы включают разработку безопасных и полезных технологий, избежание цифровой предвзятости и дискриминации, обеспечение прозрачности и объяснимости, уважение к частной жизни пользователей, ответственное использование ИИ в критически важных областях, таких как здравоохранение и безопасность; 2. Microsoft AI Ethics Guidelines. Microsoft разработала свои «Руководящие принципы по этике ИИ», согласно которым технологии должны быть понятными и доступными для пользователей, ИИ должен быть свободен от цифровой предвзятости, системы должны работать надежно и безопасно, защищая личную информацию пользователей, компании должны быть готовы нести ответственность за последствия использования ИИ; 3.

Salesforce Ethical AI. Salesforce разработала свои «Принципы этичного ИИ», в рамках которых ИИ должен служить интересам людей, технологии должны быть справедливыми и инклюзивными, пользователи должны понимать, как работают ИИ-системы, компании должны нести ответственность за последствия своих технологий.

Внутрикорпоративная проверка алгоритмов на наличие цифровой предвзятости и их корректировка — важный аспект разработки этичного искусственного интеллекта. Так, Google активно работает над выявлением и устранением цифровой предвзятости в своих алгоритмах. Компания использует внутренние инструменты для анализа данных и проверки моделей на предмет предвзятости, включая набор инструментов, которые помогают разработчикам оценивать и визуализировать предвзятость в моделях машинного обучения, позволяя проводить анализ по различным группам и выявлять потенциальные проблемы. На основе полученных результатов Google вносит изменения в алгоритмы или данные, чтобы уменьшить предвзятость. Microsoft применяет подходы к тестированию и корректировке своих алгоритмов, в частности, ее инструмент AI Fairness Checklist помогает командам проверять свои модели на наличие цифровой предвзятости на различных этапах разработки, включает в себя набор вопросов и метрик, которые помогают выявить возможные источники предвзятости. IBM использует несколько методов для проверки и корректировки своих алгоритмов, например, ее открытая библиотека AI Fairness 360 предоставляет инструменты для проверки и уменьшения цифровой предвзятости в моделях машинного обучения, включает в себя различные метрики и алгоритмы, которые помогают разработчикам анализировать свои модели. IBM проводит регулярные сессии обучения для своих разработчиков, чтобы они могли лучше понимать, как выявлять и устранять предвзятость в своих алгоритмах. Airbnb применяет методы для проверки и корректировки своих алгоритмов, компания проводит регулярные аудиты своих алгоритмов поиска и рекомендаций, чтобы выявить возможные источники цифровой предвзятости, работает над улучшением разнообразия данных, используемых для обучения моделей, минимизируя предвзятость.

Обратная связь от пользователей является важным инструментом для выявления и устранения цифровой предвзятости в системах искусственного интеллекта. Яндекс активно собирает обратную связь от пользователей через свои сервисы, такие как Яндекс.Поиск и Яндекс.Новости. Пользователи могут сообщать о проблемах с результатами поиска, что позволяет компании улучшать алгоритмы и снижать цифровую предвзятость. Сбер разработал механизмы обратной связи для своих ИИ-продуктов, таких как виртуальный ассистент Алиса. Пользователи могут сообщать о некорректных ответах или предвзятых рекомендациях, что помогает улучшать качество сервиса. Тинькофф использует обратную связь от клиентов для оценки работы своих ИИ-алгоритмов, например, в кредитовании и рекомендациях по продуктам. Клиенты могут сообщать о проблемах, что позволяет компании корректировать свои модели. МТС внедрил механизмы обратной связи для своих ИИ-сервисов, таких как чат-боты и системы рекомендаций. Пользователи могут сообщать о проблемах с цифровой предвзятостью, что помогает улучшать качество обслуживания. Google активно собирает обратную связь от пользователей через свои продукты, такие как Google Search и Google Photos. Пользователи могут сообщать о проблемах с цифровой предвзятостью, например, когда результаты поиска или алгоритмы распознавания изображений показывают неуместные или предвзятые результаты. Amazon использует отзывы пользователей для оценки качества своих рекомендаций и алгоритмов поиска. Пользователи могут оставлять отзывы о том, насколько релевантными были рекомендации, что помогает компании корректировать свои алгоритмы. Microsoft разработала платформу для сбора обратной связи от пользователей по поводу цифровой предвзятости в своих продуктах,

таких как Azure AI, что дает возможность пользователям сообщать о проблемах и получать более точные и релевантные результаты. X проводит опросы и собирает обратную связь от пользователей, чтобы улучшить свои алгоритмы модерации контента, выявляя цифровую предвзятость и улучшая качество контента на платформе.

Библиография

1. Шевцова И.В., Днепровская Н.В. Специфика производства и использования больших данных в государственном управлении // Вопросы государственного и муниципального управления. 2024. № 1 // URL: [https://vgmu.hse.ru/data/2024/03/19/2141186648/PAI_1-2024\(2\).pdf](https://vgmu.hse.ru/data/2024/03/19/2141186648/PAI_1-2024(2).pdf) (дата обращения: 12.02.2025).
2. Пилецкий И.И., Батура М.П., Шилин Л.Ю. Графовые технологии в интеллектуальной системе комплексного анализа данных Интернет-источников // Доклады БГУИР. 2020. № 5 // URL: <https://www.elibrary.ru/item.asp?edn=qnabez> (дата обращения: 12.02.2025).
3. Мирончук В. А., Иванцов К. А., Гордеев Е.С. Прогнозирование экономических циклов с использованием машинного обучения // Прогрессивная экономика. 2024. № 5. // URL: https://progressive-economy.ru/vypusk_1/prognostirovanie-ekonomicheskikh-cziklov-s-ispolzovaniem-mashinnogo-obucheniya/ (дата обращения: 12.02.2025).
4. Брычеев А.С. Применение искусственного интеллекта в органах государственной власти: вызовы и перспективы // Вестник евразийской науки. 2024. Т. 16. № 56. // URL: <https://esj.today/PDF/11FAVN624.pdf> (дата обращения: 12.02.2025).
5. Баранчук В.А. Информационная модель виртуального помощника для навигации по государственным услугам // Вестник ЮГУ. 2019. № 2 (53) // URL: <https://www.elibrary.ru/item.asp?id=41387923> (дата обращения: 12.01.2025).
6. Алексеева М.Г., Зубов А.И., Новиков М.Ю. Искусственный интеллект в медицине // МНИЖ. 2022. № 7-2 (121) // URL: <https://research-journal.org/archive/7-121-2022-july/artificial-intelligence-in-medicine> (дата обращения: 12.02.2025).
7. Шобонов Н.А., Булаева М.Н., Зиновьева С.А. Искусственный интеллект в образовании // Проблемы современного педагогического образования. 2023. № 79-4 // URL: <https://www.elibrary.ru/item.asp?edn=iprjag> (дата обращения: 12.02.2025).
8. Намиот Д.Е., Ильюшин Е.А. О киберрисках генеративного Искусственного Интеллекта // International Journal of Open Information Technologies. 2024. № 10 // URL: <http://injoit.org/index.php/j1/article/download/1977/1788> (дата обращения: 12.02.2025).
9. Апостолова Н.Н. Искусственный интеллект в судопроизводстве // Северо-Кавказский юридический вестник. 2019. № 3 // URL: https://www.researchgate.net/publication/336339778_ARTIFICIAL_INTELLIGENCE_IN_LEGAL_PROCEEDINGS (дата обращения: 12.02.2025).
10. Коданева С.И. Перспективы и риски внедрения искусственного интеллекта в государственном управлении // Социальные и гуманитарные науки. Отечественная и зарубежная литература. Сер. 4, Государство и право: Реферативный журнал. 2021. № 1 // URL: <https://www.elibrary.ru/item.asp?id=44842336> (дата обращения: 12.02.2025).
11. Орехов А.М., Чубаров Н.А. Цифровое неравенство и цифровая справедливость: социально-философские аспекты проблемы // Вестник РУДН. Серия: Философия. 2024. № 1 // URL: <file:///Users/artemkos/Downloads/38436-209800-1-PB.pdf> (дата обращения: 12.02.2025).
12. Тертышникова А.Г., Павлова У.О. Социальная эксклюзия как негативное последствие цифровизации // Теория и практика общественного развития. 2022. № 12 (178) // URL: http://teoria-practica.ru/rus/files/arhiv_zhurnala/2022/12/sociology/tertyshnikova-pavlova.pdf (дата обращения: 12.02.2025).
13. Харитонов Ю.С., Савина В.С., Паньини Ф. Предвзятость алгоритмов искусственного интеллекта: вопросы этики и права // Вестник Пермского университета. Юридические

науки. 2021. № 53 // URL: <https://www.elibrary.ru/item.asp?id=46679201> (дата обращения: 12.02.2025).

14. Понкин И.В. Государственное управление и регуляторное пространство в сфере искусственного интеллекта // Вестник Университета имени О. Е. Кутафина. 2022. № 11 (99) // URL: <https://vestnik.msal.ru/jour/article/view/1899> (дата обращения: 12.02.2025).

15. Намиот Д.Е., Ильюшин Е.А. Доверенные платформы искусственного интеллекта: сертификация и аудит // International Journal of Open Information Technologies. 2024. № 1 // URL: <https://www.elibrary.ru/item.asp?id=59852213> (дата обращения: 12.02.2025)

Результаты процедуры рецензирования статьи

Рецензия выполнена специалистами [Национального Института Научного Рецензирования](#) по заказу ООО "НБ-Медиа".

В связи с политикой двойного слепого рецензирования личность рецензента не раскрывается.

Со списком рецензентов можно ознакомиться [здесь](#).

Предметом исследования в рецензируемой публикации выступают возможности больших данных и искусственного интеллекта и связанные с их применением в государственном управлении риски.

Методология исследования базируется на обработке результатов проведенного автором анкетирования экспертного академического сообщества России в формате заполнения онлайн-таблицы с применением платформы Google в октябре–декабре 2024 г., обработке ответов 24 респондентов – ведущих российских экспертов из федеральных вузов нашей страны, представляющих все 8 федеральных округов (3 эксперта из каждого округа), обобщении сведений из научных публикаций по изучаемой теме.

Актуальность работы авторы связывают с развитием концепции больших данных, необходимостью и возможностью ее применения в сфере государственного управления. Научная новизна рецензируемого исследования, по мнению рецензента, состоит в выводах о ключевой проблеме внедрения искусственного интеллекта в сферу государственного управления – определение юридической ответственности за предвзятые цифровые решения, принимаемые с использованием искусственного интеллекта, выводы о необходимости внутрикорпоративной проверки алгоритмов на наличие цифровой предвзятости, о том, что обратная связь от пользователей является важным инструментом для выявления и устранения цифровой предвзятости в системах искусственного интеллекта.

В тексте публикации структурно выделены следующие разделы: Введение, Возможности использования больших данных и искусственного интеллекта в государственном управлении, Методология исследования, Риски использования искусственного интеллекта и больших данных в государственном управлении, Заключение и Библиография.

Авторы, исследуя возможности больших данных в современном государственном управлении, отмечают, что они могут поступать из различных источников и иметь различный формат; в связи с объемом и сложностью больших данных, постоянным процессом их производства, их эффективный анализ в современном государственном управлении практически невозможен без применения искусственного интеллекта. Отрадно, что в статье отражен зарубежный опыт применения методов анализа больших данных и искусственного интеллекта, в частности в Сингапуре, США и Китае. Сказано о том, что на основе анализа больших данных системы искусственного интеллекта могут предоставлять государственным служащим рекомендации для принятия решений, касающихся бюджета, социальных программ и других управленческих задач.

Заслуживают внимания представленные в публикации результаты экспертного опроса о рисках от внедрения искусственного интеллекта в государственное управление.

Библиографический список включает 15 источников – публикации отечественных ученых по теме статьи на русском языке, в тексте статьи имеются адресные отсылки на все источники, что подтверждает наличие апелляции к оппонентам.

Из замечаний следует отметить, что во введении уместно было бы отразить связь концепции больших данных со сферой государственного управления, а не только отразить суть и историю появления термина «большие данные».

Статья отражает результаты проведенного авторами исследования, соответствует направлению журнала «Политика и общество», содержит элементы научной новизны и практической значимости, может вызвать интерес у читателей, рекомендуется к опубликованию.