Секция «Технологии искусственного интеллекта в предоставлении государственных и муниципальных услуг»

Преодоление рисков использования искусственного интеллекта при оказании государственных услуг

Научный руководитель – Журавлев Денис Максимович

Леонов Алексей Дмитриевич

Студент (магистр)

Московский государственный университет имени М.В.Ломоносова, Высшая школа государственного администрирования (факультет), Москва, Россия

E-mail: leonov@anspa.ru

В современном мире искусственный интеллект (ИИ) революционизирует предоставление государственных услуг, обещая повышение эффективности и доступности для граждан [5]. Однако внедрение ИИ сопряжено с рядом рисков, включая вопросы конфиденциальности, точности данных и кибербезопасности. Рассмотрим некоторые риски и обозначим стратегии и методы их преодоления для обеспечения безопасности и эффективного использование ИИ в государственном секторе.

Одним из ключевых рисков является нарушение конфиденциальности и возможность утечек данных. В эпоху цифровизации государственные учреждения обрабатывают огромные объемы личной информации граждан, от персональных данных до финансовой информации. Системы ИИ, анализирующие эти данные для повышения качества и доступности услуг, могут стать мишенью для киберпреступников. Утечка данных не только нарушает права граждан на конфиденциальность, но и подрывает доверие к государственным институтам. Важно осознавать, что безопасность данных в ИИ-системах - это не единовременная задача, а постоянный процесс, требующий регулярного обновления и адаптации к новым угрозам. Для преодоления данного риска необходимо принять комплексные меры безопасности. Во-первых, критически важно применение передовых технологий шифрования данных, которые могут защитить информацию даже в случае её несанкционированного доступа. Во-вторых, должна быть внедрена многоуровневая система аутентификации и контроля доступа, чтобы минимизировать риски неавторизованного взлома. Также, регулярное проведение аудитов безопасности и непрерывное обучение персонала по кибергигиене поможет выявлять и устранять потенциальные уязвимости в ИИ-системах [2]. Важную роль играет разработка и внедрение стандартов защиты данных, которые должны учитывать специфику ИИ-технологий и быть интегрированы на всех этапах разработки и эксплуатации ИИ-систем. Кроме того, необходимо обеспечить прозрачность и контролируемость процессов ИИ, что позволит вовремя обнаруживать и исправлять ошибки, предотвращая потенциальные утечки данных. Создание правовой и нормативной базы, регулирующей использование и защиту данных в ИИ-системах, также способствует снижению рисков. Это включает в себя не только законодательные акты, но и этические нормы использования ИИ. В совокупности эти меры способствуют созданию устойчивой к угрозам среды и повышают доверие граждан к государственным услугам, основанным на ИИ.

Проблемы с законодательным регулированием являются значительным риском при использовании ИИ в государственных услугах [1]. Законы и нормативные акты часто отстают от быстрого темпа развития технологий, что создает правовой вакуум и неопределенность в вопросах ответственности, стандартов безопасности и этических норм. Например, использование ИИ для автоматизации принятия решений в государственном управлении

требует четких юридических рамок для обеспечения прозрачности, справедливости и защиты прав граждан. Отсутствие адекватного регулирования не только увеличивает риск юридических коллизий и злоупотреблений, но и подрывает доверие общества к применению ИИ в госсекторе. Для преодоления риска, связанного с отставанием законодательства от развития ИИ, необходимо активизировать работу над созданием и постоянным обновлением специализированных норм и стандартов. Это включает в себя разработку и внедрение четких юридических рамок, регулирующих использование ИИ, в том числе вопросы ответственности, безопасности данных и этических норм [6]. Важно обеспечить участие в этом процессе не только юристов и законодателей, но и специалистов в области ИИ, больших данных, кибербезопасности, а также представителей общественности и экспертного сообщества [3]. Проведение публичных консультаций и обсуждений способствует учету всех заинтересованных сторон и повышает прозрачность процесса законотворчества. Разработка международных стандартов и сотрудничество с другими странами также могут способствовать гармонизации подходов к регулированию ИИ, что особенно важно для трансграничных данных и услуг. Внедрение адаптивных нормативных актов, способных быстро реагировать на новые технологические вызовы и изменения, также является ключевой стратегией. Это обеспечит создание устойчивой и гибкой правовой базы, способной поддерживать инновации и одновременно защищать интересы граждан.

Использование искусственного интеллекта (ИИ) в предоставлении государственных услуг сопряжено с риском неточностей и предвзятости в обрабатываемых данных, что может привести к ошибочным результатам и несправедливым решениям. Поскольку ИИ-системы обучаются на основе имеющихся наборов данных, любые ошибки, неточности или предвзятость в этих данных могут привести к неправильным выводам или дискриминации при автоматизированном принятии решений [7]. Например, если данные для обучения ИИ содержат предвзятость по отношению к определенной группе людей, это может привести к систематической дискриминации при предоставлении государственных услуг, таких как социальное обеспечение, здравоохранение или образование. Таким образом, необходимо гарантировать, что данные, используемые для обучения и работы ИИ, являются точными и актуальными для всех сегментов общества. Для преодоления данного риска необходимо внедрить стратегии по улучшению качества исходных данных. Проведение тщательного аудита и очистки данных может помочь исключить ошибки и неточности. Важно также использовать методы де-идентификации данных для защиты личной информации. Для борьбы с предвзятостью необходимо обеспечить разнообразие и представительность наборов данных, включая информацию от различных групп населения. Это поможет ИИ-моделям обучаться на более широком и объективном спектре данных, снижая риск дискриминационных выводов. Применение методов машинного обучения, способных обнаруживать и корректировать предвзятость в данных, является еще одной ключевой стратегией. Разработка и внедрение этических принципов и стандартов для ИИ-систем, включая прозрачность и возможность проверки результатов, также критически важны. Повышение осведомленности и компетенций сотрудников в области этики ИИ и борьбы с предвзятостью способствует более ответственному и осознанному подходу к использованию данных [4]. В совокупности, эти меры направлены на создание справедливых и точных ИИ-систем, способствующих улучшению качества государственных услуг для всех слоев населения.
Таким образом, применение ИИ в оказании государственных услуг открывает новые возможности для повышения эффективности и доступности услуг для граждан. Однако с этими возможностями приходят и риски, такие как нарушение конфиденциальности, точности данных и кибербезопасности, требующие внимательного рассмотрения и преодоления. Как показывает обсуждение, комплексный подход к безопасности данных, законодательному регулированию и управлению качеством данных, включая стратегии борьбы

с предвзятостью и повышение прозрачности ИИ-систем, может существенно снизить эти риски. Внедрение такого подхода требует совместных усилий государственных органов, специалистов в области ИИ, законодателей и общества в целом. Развитие и адаптация законодательства, постоянное обновление технологий безопасности и этика использования ИИ станут ключом к созданию устойчивых и эффективных систем государственных услуг, способных противостоять вызовам современного цифрового мира.

Источники и литература

1) Архипов В. В., Наумов В. Б., Смирнова К. М. Пределы принятия юридически значимых решений с использованием искусственного интеллекта. Вестник Санкт-Петербургского университета. Право. 2021. Т. 12. № 4. С. 882–906.
2) Горян Э.В. Национальные подходы к применению искусственного интеллекта: опыт Сингапура. Юридические исследования. 2020. № 8. DOI: 10.25136/2409-7136.2020.8.33919. URL: https://nbpublish.com/library_read_article.php?id=33919
3) Косоруков А.А. Технологии искусственного интеллекта в современном государственном управлении. Социодинамика. 2019. № 5. DOI: 10.25136/2409-7144.2019.5.29714. URL: https://nbpublish.com/library_read_article.php?id=29714
4) Новикова И.В., Самайбекова З.К. Современные технологии стратегического управления персоналом в условиях инновационного развития предпринимательских структур. Управленческое консультирование. 2024;(1):84-95. https://doi.org/10.22394/1726-1139-2024-1-84-95
5) Талапина, Эльвира В. Искусственный интеллект и правовые экспертизы в государственном управлении. Вестник Санкт-Петербургского университета. Право 4: 865–881, 2021. https://doi.org/10.21638/spbu14.2021.404
6) Özkiziltan, D. Melanie Mitchell: Artificial intelligence—a guide for thinking humans. Genetic Programming and Evolvable Machines 23, 581–582 (2022). https://doi.org/10.1007/s10710-022-09439-7
7) Shaheen, R., & Kasi, M. Government by Algorithm: Artificial Intelligence in Federal Administrative Agencies, a Case of USA. European Journal of Technology, 5(1), 1 - 15, 2021. https://doi.org/10.47672/ejt.641