

УДК 338.246

<http://doi.org/10.35854/1998-1627-2026-3-382-398>

Методы интеллектуального анализа данных и предсказательного моделирования в задачах управленческого обеспечения национальной безопасности

Мария Евгеньевна Изотова¹✉, София Сергеевна Озеранская², Елизавета Сергеевна Борисова³

^{1, 2, 3} *Институт проблем передачи информации имени А. А. Харкевича Российской академии наук, Москва, Россия*

¹ *izotova.m@iitp.ru, <https://orcid.org/0000-0000-2761-1336>*

² *ozerskaya-sss@iitp.ru, <https://orcid.org/0009-0007-3233-7141>*

³ *borisova-es@iitp.ru, <https://orcid.org/0009-0009-2258-2404>*

Аннотация

Цель. Формирование и обоснование методического подхода к интеллектуальному анализу данных и предсказательному моделированию рисков, возникающих при применении технологий искусственного интеллекта в управленческих процессах обеспечения национальной безопасности Российской Федерации (РФ), а также выявление и систематизация механизмов их снижения на основе аналитических и прогностических инструментов.

Задачи. Выявить ключевые направления применения методов интеллектуального анализа данных и предсказательного моделирования в системе государственного управления в сфере национальной безопасности; провести системный анализ рисков и угроз, связанных с использованием ИИ-технологий в управленческих и организационных процессах; разработать типологию наиболее значимых рисков и угроз на основе аналитической обработки разнородных данных; обобщить и структурировать управленческие и технологические подходы к снижению выявленных рисков с использованием инструментов прогнозирования и поддержки принятия решений.

Методология. Методологическую основу исследования составили методы интеллектуального анализа данных, элементы предсказательного моделирования и риск-ориентированный подход, интегрированные в рамки системного анализа сложных социально-технических систем. Использованы методы качественного и сравнительного анализа стратегических и нормативно-правовых документов РФ, а также результаты отечественных и зарубежных исследований, статистические и аналитические материалы, применяемые для выявления закономерностей, оценки рисков и построения прогностических сценариев.

Результаты. На основе интеллектуального анализа данных сформирована структурированная типология рисков применения ИИ-технологий в сфере национальной безопасности, включающая в себя этико-правовые риски, киберугрозы, связанные с уязвимостью моделей и данных, риски непрозрачности и ограниченной интерпретируемости алгоритмов, зависимость качества управленческих решений от характеристик исходных данных, кадровые и технологические ограничения, а также внешние угрозы, в том числе дезинформационное воздействие и эскалацию технологической конкуренции. Показано, что использование методов предсказательного моделирования позволяет выявлять потенциальные сценарии развития таких рисков и повышать обоснованность управленческих решений.

Выводы. Результаты исследования подтверждают необходимость внедрения методов интеллектуального анализа данных и предсказательного моделирования в систему управленческого обеспечения национальной безопасности. Комплексное использование аналитических и прогностических инструментов, развитие технологического и кадрового потенциала, формирование институциональных и правовых механизмов регулирования, а также поэтапное внедрение систем контроля рисков являются ключевыми условиями повышения устойчивости и адаптивности системы национальной безопасности в процессе цифровой трансформации.

Ключевые слова: интеллектуальный анализ данных, предсказательное моделирование, искусственный интеллект (ИИ), государственное управление, национальная безопасность, технологический суверенитет

Для цитирования: Изотова М. Е., Озеранская С. С., Борисова Е. С. Методы интеллектуального анализа данных и предсказательного моделирования в задачах управленческого обеспечения национальной безопасности // *Экономика и управление*. 2026. Т. 32. № 3. С. 382–398. <http://doi.org/10.35854/1998-1627-2026-3-382-398>

Methods of data mining and predictive modeling in national security management

Mariya E. Izotova^{1✉}, Sofiya S. Ozeranskaya², Elizaveta S. Borisova³

^{1, 2, 3} *Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute), Moscow, Russia*

¹ izotova.m@iitp.ru, <https://orcid.org/0000-0000-2761-1336>

² ozerskaya-sss@iitp.ru, <https://orcid.org/0009-0007-3233-7141>

³ borisova-es@iitp.ru, <https://orcid.org/0009-0009-2258-2404>

Abstract

Aim. This work aimed to develop and substantiate a methodological approach to data mining and predictive modeling of risks arising from the application of artificial intelligence technologies in management processes ensuring national security of the Russian Federation (RF), as well as to identify and systematize mechanisms for mitigating them using analytical and predictive tools.

Objectives. We seek to identify key fields for the application of data mining and predictive modeling methods in public administration in the field of national security; conduct a systemic analysis of the risks and threats associated with the use of AI technologies in management and organizational processes; develop a typology of the most significant risks and threats based on the analytical processing of heterogeneous data; and summarize and structure management and technological approaches to mitigating identified risks using forecasting and decision-making support tools.

Methods. The study employed data mining methods, elements of predictive modeling, and a risk-based approach, integrated into the framework of systems analysis of complex socio-technical systems. Qualitative and comparative analysis of strategic and regulatory documents of the Russian Federation, as well as the results of Russian and international studies, statistical and analytical materials were used to identify patterns, assess risks, and construct predictive scenarios.

Results. Based on data mining, a structured typology of risks associated with the use of AI technologies in national security was developed. This typology includes ethical and legal risks, cyberthreats associated with model and data vulnerabilities, risks of opacity and limited interpretability of algorithms, the dependence of the quality of management decisions on the characteristics of the source data, personnel and technological limitations, and external threats, including disinformation and the escalation of technological competition. It has been revealed that the use of predictive modeling methods enables to identify potential scenarios for the development of such risks and improve the validity of management decisions.

Conclusions. The study results confirm the need to integrate data mining and predictive modeling into the national security management system. The integrated use of analytical and predictive tools, the development of technological and workforce potential, the formation of institutional and legal regulatory mechanisms, and the phased implementation of risk control systems are key factors to increasing the resilience and adaptability of the national security system during the digital transformation process.

Keywords: data mining, predictive modeling, artificial intelligence (AI), public administration, national security, technological sovereignty

For citation: Izotova M.E., Ozeranskaya S.S., Borisova E.S. Methods of data mining and predictive modeling in national security management. *Ekonomika i upravlenie = Economics and Management*. 2026;32(3): 382–398. (In Russ.). <http://doi.org/10.35854/1998-1627-2026-3-382-398>

Введение

В условиях цифровой трансформации современные государства последовательно интегрируют технологии искусственного интеллекта (ИИ) в системы государственного управления, рассматривая их прежде всего как совокупность методов интеллектуального анализа данных и предсказательного моделирования, предназначенных для повышения обоснованности и оперативности управленческих решений, качества государственных сервисов и эффективности механизмов обеспечения безопасности. В России стратегическая значимость ИИ-технологий закреплена на уровне государственных приоритетов: Национальная стратегия развития искусственного интеллекта на период до 2030 г. прямо предусматривает использование ИИ для обеспечения национальных интересов и реализации стратегических приоритетов развития¹. Аналогично и обновленная Стратегия национальной безопасности Российской Федерации (РФ) относит научно-технологическое развитие к ключевым факторам укрепления безопасности государства, предполагает применение передовых цифровых технологий, включая ИИ, для защиты информационного пространства и критически важных систем². В этой связи аргументированной и логически целостной выступает представленная в стратегических нормативно-правовых актах и официальной риторике логика российского государственно-политического руководства, в рамках которой ИИ — технология, которая способна определить будущее безопасности государства и эффективность его экономики.

В контексте этой парадигмы ИИ-технологии рассматривают как инструмент, способный существенно повлиять на эффективность функционирования социально-технических систем государственного управления, включая контуры управленческого обеспечения национальной безопасности. С одной стороны, методы интеллектуального анализа данных, машинного обучения и предсказательного моделирования расширяют возможности выявления скрытых закономерностей, прогнозирова-

ния кризисных ситуаций и поддержки принятия решений в условиях неопределенности. С другой — внедрение таких методов в автоматизированные и полуавтоматизированные системы управления формирует новые классы рисков, обусловленных сложностью алгоритмов, ограниченной интерпретируемостью моделей и высокой зависимостью качества решений от характеристик исходных данных. Иными словами, с одной стороны, источник блага, с другой — новых вызовов и угроз.

Такая логика в значительной степени определена как теоретическими разработками в области анализа данных и ИИ, так и практикой создания и внедрения ИИ-технологий в современные автоматизированные системы управления [1, с. 50]. Масштабное использование алгоритмов интеллектуального анализа данных в критически значимых управленческих контурах сопровождается не только ростом эффективности, но и появлением дополнительных угроз, способных влиять на устойчивость системы национальной безопасности. К угрозам относятся ошибки алгоритмов, накопление и распространение смещений в данных, уязвимости моделей к целенаправленным воздействиям, непредсказуемое поведение ИИ-систем в нештатных ситуациях.

Вместе с тем стремительное развитие методов машинного обучения и их внедрение в области с повышенными требованиями к надежности и безопасности, от управления транспортной инфраструктурой до поддержки решений в оборонной сфере, актуализируют задачи обеспечения безопасности, контролируемости и соответствия ИИ-систем нормативным и этическим ограничениям. Непрозрачность архитектур ряда моделей, особенно глубоких нейронных сетей, существенно затрудняет верификацию принимаемых ими решений и оценку потенциальных последствий их использования. В связи с этим для государства принципиальное значение приобретает возможность аналитического и прогностического описания рисков, связанных с функционированием ИИ-технологий в управленческих системах, поскольку неконтролируемое вне-

¹ О развитии искусственного интеллекта в Российской Федерации: указ Президента РФ от 10 октября 2019 г. № 490 (в ред. от 15.02.2024) // Президент России: офиц. сайт. URL: <http://www.kremlin.ru/acts/bank/44731> (дата обращения: 11.10.2025).

² О Стратегии национальной безопасности Российской Федерации: указ Президента РФ от 2 июля 2021 г. № 400 // Президент России: офиц. сайт. URL: <http://www.kremlin.ru/acts/bank/47046> (дата обращения: 13.10.2025).

дрение таких технологий может привести к формированию новых угроз, к которым институциональная среда оказывается не готова.

В России особую актуальность приобретает анализ влияния ИИ-технологий на систему обеспечения национальной безопасности, включая задачи военной безопасности¹. Речь идет как о применении методов интеллектуального анализа данных и предсказательного моделирования государственными органами в интересах защиты страны (в правоохранительной деятельности, разведке, кибербезопасности, управлении войсками и др.), так и о рисках, возникающих при их некорректной или преждевременной интеграции в управленческие процессы. Использование алгоритмов обработки больших данных повышает точность выявления угроз, помогает прогнозировать развитие кризисных ситуаций и координировать действия органов власти. Вместе с тем недостаточная проработанность архитектур ИИ-систем, уязвимости программно-аппаратных компонентов и ошибки моделей в критически важных контурах управления способны создавать новые точки отказа, повышать вероятность кибератак, утечек чувствительной информации и нарушений прав граждан, включая право на неприкосновенность частной жизни при массовом применении технологий биометрической идентификации [2].

Наряду с внутренними рисками, существенное значение имеют и внешние факторы, включая интенсивную конкуренцию в области автономных и интеллектуальных систем, использование ИИ-технологий недружественными государствами и негосударственными акторами, трансформацию характера угроз в киберпространстве. Отдельные субтехнологии ИИ применяли в военной сфере уже в 90-е гг. XX в. [3]. Однако в современных экспертных оценках

нередко присутствует алармизм, связанный с развитием и развертыванием военных человеко-машинных команд и автономных систем принятия решений².

В этой связи актуальной научной задачей служит изучение методов интеллектуального анализа данных и предсказательного моделирования, находящихся в основе ИИ-технологий, применяемых в государственном управлении РФ, с точки зрения их использования в управленческом обеспечении национальной безопасности. Необходимо выявить и систематизировать основные риски и ограничения, сопутствующие внедрению таких методов, разработать их типологию, оценить потенциальную значимость для устойчивости управленческих систем, определить подходы к снижению негативных эффектов. Особое внимание стоит уделить вопросам формирования государственной политики и архитектуры управления ИИ-системами, включая нормативное регулирование, организационные механизмы и международно-правовые инструменты, обеспечивающие безопасное и эффективное использование интеллектуальных технологий в интересах национальной безопасности в средне- и долгосрочной перспективе.

Методология

Для достижения поставленной цели применен комплексный исследовательский подход. В первую очередь проведен анализ нормативно-правовой базы РФ в контексте развития и использования технологий ИИ. В обзор включены стратегические документы (Стратегия национальной безопасности РФ, Национальная стратегия развития ИИ до 2030 г. и др.), федеральные законы (№ 149-ФЗ «Об информации, информационных технологиях и о защите информации»³, № 152-ФЗ «О персональных данных»⁴,

¹ Bendett S. The Role of AI in Russia's Confrontation with the West // Center for a New American Security. May 03. 2024. URL: <https://www.cnas.org/publications/reports/the-role-of-ai-in-russias-confrontation-with-the-west> (дата обращения: 02.11.2025).

² Rickli J.-M., Mantellassi F., Ladetto Q. What, Why and When? A Review of the Key Issues in the Development and Deployment of Military Human-Machine Teams // Geneva Centre for Security Policy Report. February. 2024. URL: <https://www.gcsp.ch/sites/default/files/2024-12/what-why-and-when-a-review-of-the-key-issues-in-the-development-and-deployment-of-military-human-machine-teams.pdf> (дата обращения: 02.11.2025).

³ Об информации, информационных технологиях и о защите информации: федер. закон от 27 июля 2006 г. № 149-ФЗ (последняя редакция) // Справ.-правовая система «КонсультантПлюс». URL: https://www.consultant.ru/document/cons_doc_LAW_61798/ (дата обращения: 19.10.2025).

⁴ О персональных данных: федер. закон от 27 июля 2006 г. № 152-ФЗ (последняя редакция) // Справ.-правовая система «КонсультантПлюс». URL: https://www.consultant.ru/document/cons_doc_LAW_61801/ (дата обращения: 19.10.2025).

№ 258-ФЗ «Об экспериментальных правовых режимах в сфере цифровых инноваций»¹ и др.), а также государственные программы и инициативы (национальная программа «Цифровая экономика»², федеральный проект «Искусственный интеллект»³ и т. п.). Изучение официальных документов позволило показать государственное видение возможностей и угроз, связанных с ИИ, а также действующие или предлагаемые меры регулирования.

Методологическая база исследования сформирована на основе системного анализа, методов интеллектуального анализа данных и предсказательного моделирования, применяемых к задачам управленческого обеспечения национальной безопасности. Интерпретации влияния ИИ-технологий на процессы обеспечения безопасности существенно различаются в зависимости от теоретического подхода. В научной литературе используют классические концепции (реализм, неореализм, либеральные и конструктивистские школы) и формализованные подходы, основанные на теории сложных систем, теории игр, агентном моделировании и теории управления рисками [4, с. 146–149]. В рамках настоящего исследования эти подходы рассмотрены как источники моделей и аналитических предположений, применимых к анализу поведения сложных управляемых систем в условиях неопределенности.

Эмпирическую основу исследования составляет корпус современной научной и экспертной литературы, преимущественно российских источников (монографии, статьи, аналитические доклады), посвященных применению ИИ-технологий в государственном управлении и системе обеспечения национальной безопасности. Для обработки текстовых массивов использованы методы контент-анализа, позволившие выявить и формализовать ключевые риски, ограни-

чения и предлагаемые механизмы их минимизации. Полученные результаты агрегированы в структурированные категории, пригодные для дальнейшей аналитической интерпретации и использования в моделях оценки рисков. В дополнение применен сравнительный метод, ориентированный на сопоставление российских подходов к внедрению ИИ-технологий с общемировыми тенденциями в области интеллектуального анализа данных и предсказательного управления. Сравнение проведено с учетом специфики текущей геополитической обстановки и санкционных ограничений, влияющих на доступ к технологиям, данным и международному научно-техническому взаимодействию.

В целом исследование носит преимущественно качественный аналитический характер с элементами формализации, присущими информационным наукам. Выполнена классификация рисков применения ИИ-технологий по однородным группам в зависимости от источника, характера проявления и потенциального воздействия на управленческие контуры национальной безопасности. Такая систематизация создает основу для применения методов интеллектуального анализа данных и предсказательного моделирования при оценке вероятности реализации рисков, их возможных последствий. Кроме того, для каждой выделенной группы рисков проанализированы действующие и перспективные меры государственного реагирования, рассматриваемые как элементы систем поддержки принятия управленческих решений. Это позволяет связать результаты аналитической обработки данных с практическими механизмами регулирования и повышения устойчивости управленческих систем.

Ограничительные рамки исследовательских возможностей обусловлены высокой динамикой развития ИИ-технологий и нормативно-правового регулирования в рассматриваемой сфере. Быстрое обновление алгоритмов,

¹ Об экспериментальных правовых режимах в сфере цифровых инноваций в Российской Федерации: федер. закон от 31 июля 2020 г. № 258-ФЗ (последняя редакция) // Справ.-правовая система «КонсультантПлюс». URL: https://www.consultant.ru/document/cons_doc_LAW_358738/ (дата обращения: 20.10.2025).

² Паспорт национального проекта Национальная программа «Цифровая экономика Российской Федерации»: утв. протоколом заседания президиума Совета при Президенте РФ по стратегическому развитию и национальным проектам от 4 июня 2019 г. № 7 // Гарант.ру; информ.-правовой портал. URL: <https://base.garant.ru/72296050/?ysclid=mlgf41v3vt429181750> (дата обращения: 22.10.2025).

³ Паспорт федерального проекта «Искусственный интеллект» национальной программы «Цифровая экономика Российской Федерации» (приложение № 3 к протоколу президиума Правительственной комиссии по цифровому развитию, использованию информационных технологий для улучшения качества жизни и условий ведения предпринимательской деятельности от 27.08.2020 № 17) // Справ.-правовая система «КонсультантПлюс». URL: https://www.consultant.ru/document/cons_doc_LAW_398627/?ysclid=mlgf6bi9y113947798 (дата обращения: 22.10.2025).

вычислительных архитектур и регуляторных требований приводит к тому, что полученные выводы сохраняют актуальность на момент исследования и требуют последующей актуализации по мере изменения технологической и институциональной среды.

Результаты

На базе проведенного анализа можно выделить несколько ключевых *категорий рисков*, связанных с внедрением технологий ИИ в деятельность органов государственной власти и обеспечения безопасности РФ.

Технологические риски и уязвимости

К этой категории отнесены проблемы, вытекающие из природы высокотехнологичных систем ИИ, а также из их способности имитировать человеческое мышление [5, р. 29–53]. Во-первых, это риск ошибок и сбоев алгоритмов. ИИ-системы (например, системы поддержки принятия решений) могут давать сбои или неверные выводы вследствие недостатков обучающих данных, неожиданных сценариев или так называемой предвзятости алгоритмов (*bias*). Как пишет известный американский теоретик Дж. Най (мл.), в таких критически важных областях, как управление транспортом, энергетикой, вооруженными силами, ошибка ИИ может привести к серьезным инцидентам¹. Во-вторых, уязвимость перед кибератаками и внешним воздействием: сложные алгоритмы могут быть подвержены атакам типа «отравление данных» (*data poisoning*) или эксплуатации уязвимостей в программном коде, что способно вывести из строя государственные системы или, что еще хуже, повернуть их против пользователей. Все это делает проблемы повышения результативности выявления угрозы кибератак и связанных с ними рисков одними из наиболее актуальными и масштабными в контексте безопасности применения ИИ-технологий в управленческом обеспечении национальной безопасности РФ [6].

Рост количества компьютерных атак на российские информационные ресурсы признан одной из основных угроз национальной безопасности. Использование ИИ одновременно порождает новые векторы атак, но и может служить средством их отра-

жения. Наконец, технологические риски включают в себя проблему непрозрачности (черного ящика) алгоритмов: многие современные модели (особенно нейросетевые) трудно объяснимы, что затрудняет аудит их правильности и надежности [7]. Без специальных мер (в их числе — разработка объяснимого ИИ, протоколирование решений) государственные органы могут оказаться в ситуации, при которой не смогут обосновать принятое автоматизированной системой решение или обнаружить причину сбоя.

Риски информационной безопасности и конфиденциальности

Государственные информационные системы оперируют большими объемами данных, в том числе персональных и секретных. Внедрение ИИ увеличивает потенциал и одновременно угрозы в этой сфере. С одной стороны, ИИ помогает автоматизировать мониторинг киберинцидентов, выявлять аномалии и фейковые сообщения (например, нейросети используют для распознавания дезинформации). С другой — возникают новые типы угроз информационной безопасности, которые не характерны для прежних ИТ-систем. Например, использование облачных ИИ-сервисов может приводить к передаче данных сторонним организациям; генеративные модели могут по косвенным признакам восстанавливать исходные персональные данные из обезличенных наборов; широкая доступность алгоритмов упрощает злоумышленникам создание реалистичных фальшивых контентов (*deepfakes*) для информационных атак, в том числе позволяет государствам и военно-политическим альянсам использовать дипфейки в гибридных войнах [8; 9, с. 64].

Стратегия национальной безопасности РФ прямо предупреждает, что использование иностранного ИТ-оборудования и программного обеспечения повышает уязвимость отечественной инфраструктуры. Это особенно актуально для ИИ-систем, обучение и функционирование которых часто зависят от иностранных платформ и компонентов (п. 55). Кроме того, конфиденциальность данных, принадлежащих гражданам, находится под угрозой, если алгоритмы массово собирают и обрабатывают личные сведения.

¹ Nye J. S., jr. AI and National Security // Project Syndicate. Jule 31. 2024. URL: <https://www.project-syndicate.org/commentary/ai-national-security-some-benefits-and-many-risks-by-joseph-s-nye-2024-07> (дата обращения: 24.10.2025).

Без должного контроля возможны несанкционированная утечка или использование таких данных. Этот риск государством осознан: требования защиты персональных данных при разработке и обучении моделей ИИ теперь отнесены к числу важных условий развития отрасли.

Правовые и регуляторные риски

Правовая система часто не успевает за технологическими инновациями. В случае с ИИ это выражено в регуляторных пробелах и барьерах, которые одновременно мешают эффективному внедрению технологий и не обеспечивают достаточных гарантий от злоупотреблений [10]. Проблема правовой неопределенности проявляется, например, в вопросе ответственности, в частности о том, кто будет нести ответственность за ущерб, причиненный решением, принятым системой ИИ: разработчик алгоритма, государственный служащий, утвердивший решение, или государство. Такая постановка вопроса актуальна и для зарубежных систем государственного управления [11]. В действующем законодательстве отсутствуют прямые нормы о правосубъектности или ответственности ИИ-систем, что создает риск правового вакуума при возможных инцидентах.

Анализ этой группы рисков выявляет еще один аспект: недостаточная адаптированность нормативных актов под ИИ. Многие законы и инструкции предполагают, что решения принимают люди, и применение автоматизированных систем формально может противоречить устаревшим нормам. К примеру, требования к документированию решений, процедурам согласования и т. п. могут не учитывать участие алгоритмов. Но существуют избыточные барьеры: бюрократические процедуры согласования инноваций, отсутствие стандартов. Все это тормозит внедрение ИИ в государственном управлении.

Этические и социальные риски

К этой группе следует отнести потенциально негативные последствия применения ИИ для общества, общественных ценностей и прав граждан [12]. Этические проблемы ИИ, активно обсуждаемые в мире, в том числе в России, нашли наиболее полное отражение в фундаментальном международном правовом

акте — Этических рекомендациях ЮНЕСКО¹. В публичном управлении особенно остро стоят вопросы недискриминации и справедливости алгоритмов. Известны случаи, в которых алгоритмы распознавания лиц проявляли расовые или гендерные смещения в точности работы, либо системы отбора персонала отсекали кандидатов по формальному признаку, переняв предвзятость из обучающих данных.

В контексте государственного управления такое непреднамеренное воспроизведение несправедливости недопустимо, поскольку подрывает принципы равенства перед законом. Еще одна этическая дилемма — прозрачность принимаемых решений. Граждане и организации должны иметь возможность понять основание, по которому, например, им отказано в государственной услуге, если решение принято с участием ИИ. Принцип прозрачности и объяснимости алгоритмов закреплен среди главных в обновленной Стратегии развития ИИ, что говорит о понимании этой проблемы. Возможен и риск нарушения приватности: массовое применение технологий распознавания лиц, голозов, анализ социальных сетей государством ради безопасности может войти в противоречие с правом на частную жизнь и другими гражданскими свободами [13]. Если меры безопасности реализованы непрозрачно, это чревато снижением доверия общества к государству.

Общество может негативно отреагировать на ошибки ИИ (например, ошибочная санкция или арест на основе алгоритма прогноза правонарушений). Такие инциденты способны вызвать общественный резонанс и протесты. В этом аспекте этические риски напрямую связаны с социально-политическими: ненадлежащее использование ИИ может подорвать легитимность действий властей. Поэтому государство должно демонстрировать ответственное отношение. В частности, Кодекс этики в сфере ИИ (принят в РФ в 2021 г.) гласит о том, что технологии должны быть человекоориентированными, не подменять человека в критически важных решениях, а пользователи должны иметь право отказаться от общения с ИИ и требовать пересмотра решения человеком. Соблюдение указанных принципов — важная составляющая снижения социальных рисков.

¹ Recommendation on the ethics of artificial intelligence // UNESCO. November 23. 2022. URL: <https://unesdoc.unesco.org/ark:/48223/pf0000381137> (дата обращения: 01.11.2025).

Внедрение ИИ в работу органов государственной власти — это не только технический, но и организационный вызов. Дефицит компетентных кадров способен стать «узким местом»: для разработки, внедрения и эксплуатации ИИ-систем требуются специалисты (разработчики, аналитики данных, ИТ-архитекторы и так называемые кибербезопасники), которых в штате традиционных государственных органов может не быть в достаточном количестве. Статистика подтверждает наличие этой проблемы: в России планируется увеличить количество выпускников вузов по ИИ-направлениям с 3 до 15,5 тыс. в год к 2030 г., что отражает текущий дефицит компетенций¹. Кроме того, сопротивление организационной культуре изменениям может тормозить цифровизацию. Государственный аппарат отличается сложившимися бюрократическими процедурами, и процесс внедрения новых технологий часто сталкивается с инерцией, опасениями сотрудников (например, это может быть страх потерять работу вследствие автоматизации) и недостаточной гибкостью. Как следствие, низкий уровень внедрения ИИ в государственном управлении — сам по себе риск, поскольку Россия рискует отстать в эффективности государственного менеджмента. Обновленная Стратегия развития ИИ прямо указывает на недостаточно высокую степень внедрения ИИ в деятельность органов власти как на один из вызовов, сформировавшихся в 2022–2023 гг. Это означает, что, несмотря на провозглашенные цели цифровизации, практическое использование ИИ пока представляется ограниченным. Уместно дополнить к изложенному и риск межведомственной разрозненности: при отсутствии единого стандарта и координирующего центра разные ведомства могут развивать ИИ несогласованно, дублируя усилия или создавая системы, не совместимые друг с другом. Таким образом, организационные риски требуют не только подготовки кадров, но и изменений в управленческих процессах, создания новых структур (например, центров компетенций по ИИ при органах власти) и налаживания межведомственного взаимодействия.

¹ В Минобрнауки рассказали о планах по числу выпускаемых специалистов в сфере ИИ // ТАСС. 2023. 23 ноября. URL: <https://tass.ru/ekonomika/19361221> (дата обращения: 12.10.2025).

Эту группу составляют риски, связанные с технологической независимостью и внешними факторами. В условиях обострившейся международной обстановки доступ РФ к ряду передовых технологий ограничен односторонними ограничительными мерами недружественных государств. В сфере ИИ это проявляется в нехватке высокопроизводительных процессоров и оборудования для обучения моделей, ограничении доступа к некоторым программным продуктам и исследовательским сотрудничествам. Зависимость от импорта технологий ИИ рассматривается как угроза: в Стратегии развития ИИ сделан акцент на «нехватке вычислительных мощностей» и «недостаточном развитии отечественных решений в области ИИ». Если государственные органы будут критически зависеть от иностранного программного обеспечения или облачных платформ, возникает риск, что в случае санкций или конфликта эти сервисы станут недоступны, парализуя важные системы. Более того, использование чужих технологий может нести скрытые угрозы, в том числе закладки, шпионские функции, сбор данных. Такой сценарий упоминается как «повышение уязвимости ИТ-ресурсов при использовании иностранных ИКТ». Поэтому технологический суверенитет в сфере ИИ — это не просто вопрос престижа, а элемент национальной безопасности. Российская политика в последние годы реагирует на этот риск, провозглашая приоритет отечественных разработок, полной локализации критических технологий и импортозамещения в государственных структурах.

В дополнение к зависимости геополитический риск включает в себя опасность международной гонки ИИ-вооружений: ряд стран разрабатывают автономные системы оружия, и отсутствие международных норм может привести к неконтролируемому распространению таких систем. Это создает угрозу стратегической стабильности. Россия на международной арене выступает за многостороннее регулирование ИИ под эгидой ООН, противопоставляя его фрагментарным западным инициативам. В данном контексте продвижение собственной модели глобального



Рис. 1. Формирование замкнутого контура управления системой принятия решений в сфере обеспечения национальной безопасности с использованием ИИ

Fig. 1. Formation of a closed-loop control of system for national security decision-making using AI

Источник: составлено авторами.

управления ИИ тоже можно считать мерой снижения стратегических рисков.

Оценка значимости рисков

Изложенные выше категории рисков различаются по своей природе, однако они взаимосвязаны и в совокупности определяют проблемное поле внедрения ИИ в государственное управление. С точки зрения национальной безопасности наиболее критичны риски, которые могут привести к прямому ущербу безопасности государства и граждан. К ним относятся технологические и киберриски (успешная кибератака на ключевую ИИ-систему, например МВД России или Минобороны России, имела бы тяжелые последствия). Не менее значим стратегический риск потери технологического суверенитета, то есть зависимость от зарубежных ИИ-решений в условиях геополитического противостояния может привести к утрате контроля над информационной безопасностью и экономике. Правовые и этико-социальные риски имеют более опосредованное действие, но тоже существенны: правовой

вакуум способен замедлить развитие (это чревато технологическим отставанием, и тоже фактор риска), а этические проблемы могут ослабить общественную стабильность внутри страны. С учетом этого эффективной мерой преодоления комплекса иных угроз представляется формирование замкнутого контура управления системой принятия решений в сфере обеспечения национальной безопасности с применением и на основе ИИ-технологий, как видно на рисунке 1.

Предложенная авторами схема на рисунке 1 показывает концептуальную военно-доктринальную модель замкнутого контура управления применением ИИ-технологий в задачах управленческого обеспечения национальной безопасности РФ. Верхний уровень схемы формируют нормативный контур, аналитическое ядро ИИ и технологическая база, которые в совокупности задают регуляторные ограничения, аналитические механизмы интеллектуального анализа данных и предсказательного моделирования, а также материально-техническую основу функционирования систем поддержки

принятия решений. Нижний уровень отражает контуры выработки и реализации управленческих решений в государственных и силовых структурах. Обратная связь, замыкающая контур, обеспечивает адаптивную корректировку нормативных и аналитических компонентов на основе результатов практического применения ИИ-технологий, что позволяет поддерживать устойчивость и управляемость системы в условиях изменяющейся обстановки и динамики угроз.

Итак, несмотря на то, что потенциал негативных последствий высок во всех группах (от точечных сбоев и утечек данных до масштабных кризисов доверия или безопасности), тем не менее государственная политика может ранжировать риски по вероятности и тяжести. Видимо, наивысший приоритет получают угрозы безопасности (кибератаки, внешнее давление), за ними — риски, влияющие на устойчивость системы (законодательные проблемы, кадровый голод), а затем — долгосрочные этические и социальные вопросы. Далее в статье нами рассмотрены меры, которые предпринимают или которые могут быть предложены для минимизации указанных рисков.

Обсуждение

Меры государственной политики по минимизации рисков ИИ

Анализ показал, что в России сформировано понимание перечисленных проблем, и на государственном уровне предпринимают меры для их решения. Тем не менее масштабы вызова требуют комплексной и опережающей политики. Далее представлены основные направления и инструменты, которые могут обеспечить безопасную интеграцию ИИ в государственное управление с акцентом на национальную безопасность.

Совершенствование нормативно-правового регулирования

Законодательная база адаптируется к эпохе ИИ, хотя остается много открытых вопросов, в том числе в сфере цифровой экономики [14]. В 2020 г. принят Федеральный закон № 258-ФЗ «Об экспериментальных правовых режимах в сфере цифровых ин-

новаций», вводящий систему так называемых регуляторных песочниц. Благодаря им отраслевые регуляторы могут временно смягчать требования для пилотных проектов с ИИ, чтобы последние проходили обкатку без риска нарушения закона. Изложенный подход позволяет одновременно стимулировать инновации и собирать информацию для последующей доработки законодательства. В настоящее время сфера применения экспериментальных правовых режимов расширяется, в частности предполагается запуск песочниц именно в государственном управлении. Минцифры России в 2025 г. подготовило проект постановления Правительства РФ о проведении эксперимента по использованию генеративного ИИ в государственном управлении. В рамках этого пилотного проекта планируется опробовать типовые ИИ-сервисы для ведомств (по направлениям: управление данными, обработка текстовых обращений, поддержка кадровых решений, анализ рисков, контент-менеджмент)¹.

В правилах эксперимента отдельно указаны меры безопасности: запрещено использовать генеративный ИИ для обработки государственной тайны, а также для прогнозирования социально-экономических процессов, что должно предотвратить утечку чувствительной информации и некорректное влияние ИИ на решение политических задач. Кроме того, использовать будут преимущественно отечественные или открытые программные решения, что снизит риск зависимости от зарубежных сервисов. Ожидается, что результаты эксперимента помогут определить необходимые изменения в регламентах работы органов власти и технические требования к ИИ-системам.

В более широком плане Минцифры в 2025 г. представило проект концепции регулирования ИИ до 2030 г., который означает переход от точечных инициатив к системной правовой базе. В концепции провозглашены новые принципы: человекоориентированный подход, доверенность (прозрачность и контроль алгоритмов), технологический суверенитет и др.

Можно предположить, что закрепление этих принципов в законодательстве — следующий шаг. Вероятно, будет необходим отдельный федеральный закон об обращении

¹ Проект постановления Правительства РФ «О проведении эксперимента по использованию генеративного искусственного интеллекта в государственном управлении»: подготовлен Минцифры России 23 июля 2025 г. // Гарант.ру: информ.-правовой портал. URL: <https://www.garant.ru/products/ipo/prime/doc/56928658> (дата обращения: 28.10.2025).

ИИ-систем (как это делают в ряде стран), с учетом которого в том числе будут решены вопросы сертификации ИИ для критических применений, распределения ответственности, правового статуса решений, принятых ИИ. На переходный период разумно использовать подзаконные акты и стандарты: например, установить при Правительстве РФ межведомственную комиссию по ИИ, которая будет оперативно разрабатывать рекомендации и проекты поправок. Сегодня фундаментом служат такие акты, как Указ Президента РФ от 10 октября 2019 г. № 490 «О развитии искусственного интеллекта в Российской Федерации» (речь идет о Национальной стратегии развития ИИ на период до 2030 г.); определены общие правовые аспекты развития ИИ, Правительству РФ поручено выработать требования к технологиям. В исполнение этих поручений государственные органы разрабатывают соответствующие методические рекомендации. Постепенное устранение пробелов в регулировании и обновление норм под реалии ИИ позволит сократить неопределенность и повысить предсказуемость для участников.

Обеспечение технологического суверенитета и безопасности инфраструктуры

Как указано ранее в статье, зависимость от импортных технологий рассматривается как неприемлемый риск. Поэтому государственная политика ориентирована на развитие отечественной ИИ-экосистемы. В Национальной стратегии развития ИИ на период до 2030 г. отражено, что в условиях санкций главной целью стало создание максимально независимой национальной среды ИИ. Практические шаги предусматривают поддержку отечественных разработчиков программного обеспечения и алгоритмов (гранты, государственные закупки, льготные режимы для ИТ-компаний), стимулирование производства элементной базы, то есть собственных чипов, серверов для обучения моделей (в 2023 г. запущены инициативы по производству отечественных графических процессоров), импортозамещение в действующих системах государственных

органов (переход на российские операционные системы, системы управления базами данных и облака). Например, поставлена задача к определенному сроку в полной мере перевести критическую инфраструктуру и органы власти на российское программное обеспечение и оборудование, что находит отражение и в законопроектах на федеральном уровне¹.

Помимо этого, уделяют внимание киберустойчивости ИИ-систем: речь идет о разработке «доверенных» ИИ-моделей, прошедших проверку на отсутствие уязвимостей и закладок, создании национального репозитория безопасных решений ИИ. Такая работа координируется, в частности, профильными центрами при Совете безопасности РФ и ФСБ России, отвечающими за техническую защиту информации. Доктрина информационной безопасности РФ² и Стратегия национальной безопасности РФ требуют постоянного совершенствования методов защиты с учетом передовых технологий. Сегодня это означает использование ИИ для кибербезопасности (системы выявления вторжений на базе ИИ и т. д.) и против злонамеренных ИИ (разработка методов противодействия генеративным фейкам, фильтрации контента). В целом курс на технологический суверенитет приносит результаты: так, по заявлениям официальных лиц, в России созданы собственные фреймворки для машинного обучения, тренируются большие языковые модели, запускают пилотные зоны 5G без импортного оборудования и т. п. Это снижает долгосрочные риски, хотя требует значительных инвестиций и времени.

Развитие кадрового потенциала и научных исследований

Без квалифицированных специалистов и фундаментальной научной базы успешная интеграция ИИ невозможна. Поэтому одной из приоритетных мер служит подготовка кадров для работы с ИИ в государственном секторе в контексте обеспечения национальной безопасности. В течение последних лет запущены образовательные программы

¹ О внесении изменений в Федеральный закон «О безопасности критической информационной инфраструктуры Российской Федерации»: проект Федерального закона № 581689-8 (ред., внесенная в ГД ФС РФ, текст по состоянию на 21 марта 2024 г.) // Справ.-правовая система «КонсультантПлюс». URL: <https://www.consultant.ru/cons/cgi/online.cgi?req=doc&base=PRJ&n=245725&ds> (дата обращения: 29.10.2025).

² Об утверждении Доктрины информационной безопасности Российской Федерации: указ Президента РФ от 05 декабря 2016 г. № 646 // Президент России: офиц. сайт. URL: <http://www.kremlin.ru/acts/bank/41460> (дата обращения: 25.10.2025).

по ИИ во многих вузах, создают магистратуры и кафедры (например, в РАНХиГС при Президенте РФ, МГУ, МФТИ, ИТМО и др.), вводят специальные курсы для государственных служащих по цифровым навыкам. Согласно федеральным целевым показателям, количество выпускников в сфере ИИ должно значительно увеличиться. Важно не только количество, но и качество; для национальной безопасности нужны элитные специалисты, способные разрабатывать конкурентоспособные алгоритмы (это обеспечивается через поддержку олимпиад, конкурсов, молодежных лабораторий по ИИ). Одновременно на государственном уровне поддерживают научные исследования в области ИИ и анализа рисков. Российский научный фонд, Российский фонд фундаментальных исследований финансируют исследования, в том числе по темам этики ИИ, надежности нейросетей, методам объяснимости. Создают в вузах нашей страны научные центры, такие как Институт перспективных исследований проблем искусственного интеллекта и интеллектуальных систем при МГУ имени М. В. Ломоносова¹, в котором тоже изучают вопросы безопасности и эффективности ИИ. Накопление научных знаний позволит лучше понимать и предсказывать поведение сложных ИИ-систем, разрабатывать методы их тестирования. К тому же в обновленной Стратегии развития ИИ предусмотрена задача анализа международного опыта обеспечения безопасности ИИ, то есть ученые и эксперты должны мониторить лучшие практики, существующие в мире, и адаптировать их для РФ.

Этическое регулирование, повышение доверия

Помимо сугубо правовых норм, важную роль выполняют саморегуляция отрасли и формирование этических стандартов [15]. В 2021 г. на площадке Альянса в сфере ИИ² (в который входят ведущие российские ИТ-корпорации и государственные организации) принят Кодекс этики ИИ³. Он заложил основы ответственного отношения к ИИ: принципы приоритета человека,

безопасности, конфиденциальности, непричинения вреда, разделения ответственности и др. Более 360 организаций присоединились к этому Кодексу, и впоследствии разработаны отраслевые приложения (например, кодексы этики ИИ в медицине, финансах); всего в России разработано более 100 национальных и 19 международных стандартов в сфере регулирования и развития ИИ⁴.

Государство поддерживает эти инициативы: Аналитический центр при Правительстве РФ и Банк России участвовали в разработке указанного Кодекса, а при Минэкономразвития России создана комиссия по его реализации. Подобные мягкие механизмы регулирования дополняют закон, способствуя формированию культуры ответственного внедрения ИИ. В кодексах даны практические рекомендации: например, как указано ранее, дать гражданину возможность требовать пересмотра решения ИИ человеком, обеспечивать объяснимость алгоритмов, не использовать ИИ для недобросовестных или незаконных целей. Соблюдение этих норм будет влиять на репутацию организаций.

Для государственного сектора целесообразно разработать аналогичный этический кодекс использования ИИ в публичном управлении, учитывающий специфику государственной службы и права граждан. Кроме того, важно вести разъяснительную работу с обществом: информировать о целях и способах применения ИИ государственными органами, о мерах защиты прав граждан. Если население видит, что государство ответственно относится к новым технологиям, это повышает доверие. Вовлечение гражданского общества, например обсуждение проектов ИИ (как в случае с экспериментом Минцифры России, который вынесен на общественное обсуждение), также помогает учитывать возможные этические аспекты и получать обратную связь. В итоге сочетание законодательных требований и добровольных этических обязательств формирует среду доверия, в которой ИИ воспринимается не как угроза, а как инструмент, действующий на благо общества. Уровень

¹ Наши лаборатории // Институт перспективных исследований проблем искусственного интеллекта и интеллектуальных систем МГУ имени М. В. Ломоносова. URL: <https://www.iai.msu.ru/> (дата обращения: 20.10.2025).

² Миссия // Альянс в сфере искусственного интеллекта. URL: <https://a-ai.ru> (дата обращения: 22.10.2025).

³ Кодекс этики в сфере искусственного интеллекта (от 26 октября 2021 г.) // Гарант.ру: информ.-правовой портал. URL: <https://base.garant.ru/406862712> (дата обращения: 16.10.2025).

⁴ К кодексу этики ИИ присоединились уже более 360 организаций и органов власти РФ // ТАСС. 2024. 17 апреля. URL: <https://tass.ru/ekonomika/20569549> (дата обращения: 29.10.2025).

доверия планируется существенно повысить. Для этого будет реализован комплекс мероприятий: просветительские кампании, показы успешных кейсов ИИ без негативных последствий, открытые тестирования и др.

Организационные меры: координация и управление рисками

Для эффективного контроля над рисками государству необходимы соответствующие институты и процессы. Одно из направлений — создание структур, ответственных за координацию политики в области ИИ. В настоящее время функции распределены между разными органами (Минцифры России — развитие ИИ и цифровой экономики, Минобороны России и спецслужбы — военное применение, Минэкономразвития России — вопросы регулирования и этики, Минобрнауки России — наука и кадры, при Правительстве РФ действует экспертный совет по ИИ¹ и т. д.). Востребованным, на наш взгляд, было бы учреждение единого центра, в частности создание Агентства (или Комитета) по ИИ, которое консолидировало бы работу, особенно в аспекте обеспечения безопасности и выработки единых стандартов для ведомств. Пока эту роль частично выполняют профильный проектный офис и комитеты АНО «Цифровая экономика». Далее процесс управления рисками должен стать неотъемлемой частью внедрения ИИ.

Для проектов с участием ИИ на уровне государственных органов целесообразно внедрять процедуры оценки рисков на стадии планирования: ответственные лица должны заранее идентифицировать потенциальные риски (этические, правовые, киберриски) и прописывать меры по их снижению. Например, при внедрении системы видеоаналитики в городе следует оценить риск нарушения прав частной жизни граждан или функционирования организаций, сразу предусмотреть технические ограничения (размывание лиц в нецелевых задачах и т. д.), юридические гарантии (правила доступа к данным) и коммуникационную стратегию с населением. В ряде стран подобные процедуры становятся обязательными. Россия движется в этом направлении. Так, упомянутая концепция регулирования

предусматривает риск-ориентированный подход.

Необходимы тестирование и сертификация важных ИИ-систем. Возможным решением было бы создание испытательного центра (лаборатории) для проверки алгоритмов, которые планируются к использованию в критической инфраструктуре или сфере безопасности, на предмет стойкости к атакам, соответствия требованиям по защищенности и точности. Еще одна организационная мера — развитие партнерств между государством и частными компаниями-разработчиками ИИ. Многие компетенции сконцентрированы в компаниях (среди них — Сбер, Яндекс, Ростелеком, Крок), и государство может привлекать их к совместным проектам, что снижает риск дефицита экспертизы. При этом государственный сектор должен вырабатывать типовые решения и платформы, доступные для всех органов власти (сегодня, например, создают Единую информационную платформу², в частности национальную систему управления данными (НСУД), к которой будут подключать аналитические сервисы с элементами ИИ). Стандартизованные решения проще защищать и контролировать, в отличие от множества разрозненных.

Селективное международное сотрудничество и глобальные инициативы

Хотя в условиях санкций международное взаимодействие РФ затруднено, в долгосрочной перспективе участие в глобальном регулировании ИИ отвечает интересам национальной безопасности. Россия выступает за разработку международных норм, запрещающих наиболее опасные формы использования ИИ (например, полностью автономное летальное оружие) и обеспечивающих справедливые условия технологического обмена. В ООН российская сторона продвигает идею многостороннего договора по ИИ, обсуждает эти вопросы на площадках БРИКС, ШОС. Сотрудничество с дружественными странами (Китаем, Индией, странами СНГ) в области ИИ тоже видится значимым: совместные исследования, обмен данными для обучения моделей, создание альтернативных международных инфраструктур (например, аналог

¹ Национальный центр развития искусственного интеллекта при Правительстве Российской Федерации (НЦРИИ) // Национальный исследовательский университет «Высшая школа экономики». URL: <https://aicentre.hse.ru/?ysclid=mlgiein39861051987> (дата обращения: 02.11.2025).

² Что можно делать в ЕИП НСУД // Единая информационная платформа национальной системы управления данными (ЕИП НСУД). URL: nsud.gosuslugi.ru (дата обращения: 01.11.2025).

SWIFT для данных ИИ) помогают снизить зависимость от монополии нескольких ведущих держав и корпораций. Кроме того, международный диалог по этике и безопасности ИИ позволяет перенимать лучшие практики. В этой сфере заметны усилия ЮНЕСКО. Так, в 2021 г. приняты Рекомендации ЮНЕСКО по этике в сфере ИИ, в разработке которых наша страна активно участвовала и к которым РФ присоединилась¹.

Реализация Рекомендаций ЮНЕСКО на национальном уровне будет еще одним шагом к укреплению принципов ответственного использования ИИ. Таким образом, интернационализация проблемы и участие в формировании «правил игры» на глобальной арене рассматривают как часть стратегии снижения рисков: это поможет избежать неконтролируемой гонки ИИ и обеспечить учет национальных интересов РФ в будущей мировой архитектуре управления ИИ.

В совокупности перечисленные направления образуют каркас государственной политики, направленной на то, чтобы максимизировать пользу от ИИ и минимизировать связанные угрозы, как следует из таблицы 1. Достигнутые результаты (принятие стратегии и этического кодекса, изменение законов, запуск пилотных проектов) свидетельствуют о прагматичном подходе, то есть государство стремится опережать возможные негативные сценарии, а не реагировать постфактум. Тем не менее многое предстоит сделать: от разработки конкретных технических стандартов и методик оценки рисков до формирования в обществе доверия к ИИ как инструменту повышения благосостояния и безопасности.

Выводы

Проведенное исследование подтверждает, что внедрение технологий ИИ, основанных на методах интеллектуального анализа данных и предсказательного моделирования, в систему государственного управления и управленческого обеспечения национальной безопасности РФ формирует качественно новые контуры принятия решений, одновременно повышая эффективность управления и усиливая чувствительность системы к рискам технологического и институционального характера. Использование ИИ-технологий в критически значимых управленческих процессах требует перехода

от фрагментарного регулирования к централизованной и формализованной модели управления рисками, интегрированной в архитектуру национальной безопасности.

Результаты анализа показывают, что наибольшая концентрация угроз связана с техническими и алгоритмическими особенностями ИИ-систем, включая уязвимости моделей, деградацию качества решений при изменении характеристик данных, а также возможность целенаправленных воздействий на алгоритмы и информационные потоки. Существенное значение имеют риски кибербезопасности и утечек чувствительной информации, возникающие при интеграции ИИ в распределенные информационные и управленческие системы. Дополнительную нагрузку на устойчивость управленческих контуров создают нормативно-правовые ограничения, дефицит квалифицированных кадров и сохраняющаяся зависимость от зарубежных аппаратных и программных компонентов, а также внешние факторы, связанные с ускоренной технологической конкуренцией и применением ИИ-технологий недружественными странами и негосударственными акторами.

В ходе исследования выявлено, что наиболее критичными являются риски относительно базовых элементов национальной безопасности, включая киберустойчивость, технологический суверенитет и защиту прав граждан. Вместе с тем риски второго порядка, такие как институциональная инерция, фрагментарность регулирования и снижение доверия к автоматизированным управленческим решениям, при накопительном эффекте способны трансформироваться в системные угрозы управляемости. В этой связи управление рисками применения ИИ-технологий целесообразно рассматривать как обязательный компонент государственной политики в сфере национальной безопасности, а не как вспомогательное направление цифровой трансформации.

Выполненный анализ свидетельствует о необходимости формирования замкнутого контура управления ИИ-технологиями, ориентированного на обеспечение контролируемости, верифицируемости и предсказуемости алгоритмов, используемых в управленческих и силовых контурах. Ключевая роль в этой модели отведена развитию отечественной вычислительной инфраструктуры,

¹ Recommendation on the ethics of artificial intelligence // UNESCO. November 23. 2022. URL: <https://unesdoc.unesco.org/ark:/48223/pf0000381137> (дата обращения: 01.11.2025).

Базовые риски внедрения ИИ-технологий и механизмы их минимизации

Table 1. Basic risks of implementing AI technologies and mechanisms for their mitigation

Тип риска	Описание	Механизмы минимизации
Этические и правовые проблемы	Алгоритмическая предвзятость, автономность ИИ-систем, утечка персональных данных	Разработка и внедрение национальных этических кодексов ИИ. Уточнение правового статуса и зоны ответственности. Совершенствование законодательства о защите персональных данных
Киберугрозы и уязвимости ИИ-систем	Манипуляция входными данными, вредоносное вмешательство в модели	Создание стандартов кибербезопасности для ИИ. Аудит и сертификация ИИ-систем. Использование защищенных сред обучения
Проблема черного ящика (непрозрачность)	Отсутствие возможности интерпретировать решения ИИ	Развитие так называемого интерпретируемого ИИ. Внедрение требований к прозрачности алгоритмов в регуляторную практику
Зависимость от качества данных	Искажение результатов при обучении на ошибочных или предвзятых данных	Создание единых стандартов для датасетов. Государственные хранилища верифицированных данных
Кадровый дефицит	Недостаток специалистов с компетенциями в области ИИ и информационной безопасности	Поддержка образовательных программ по ИИ и безопасности. Формирование кадрового резерва для государственной службы. Стимулирование НИОКР
Технологическая зависимость от импорта	Уязвимость критических систем ввиду использования иностранных решений	Импортозамещение в сфере ИИ. Поддержка отечественных разработчиков. Формирование национальной ИИ-инфраструктуры
Внешние угрозы (киберпреступность, терроризм, дезинформация)	Злоупотребление ИИ со стороны негосударственных акторов (в том числе при поддержке государственных акторов)	Создание специальных подразделений в органах безопасности. Повышение устойчивости инфраструктуры к дезинформации. Международное сотрудничество в сфере киберугроз (селективное)
Гонка ИИ-технологий в военной сфере	Эскалация нестабильности и стратегических рисков	Разработка международных механизмов контроля. Допустимая прозрачность военных ИИ-программ. Повышение «человеческого» контроля автономными системами вооружений, персонализация ответственности при принятии решений

Источник: составлено авторами.

программных платформ и средств защиты данных, что снижает технологические и геополитические риски и создает основу для независимого масштабирования ИИ-решений. Интеграция методов интеллектуального анализа данных и предсказательного моделирования в системы поддержки принятия решений позволяет обеспечить переход от реактивного реагирования на угрозы к превентивному управлению, основанному на сценарном анализе и оценке вероятности реализации негативных событий.

Вместе с тем полученные результаты говорят о том, что отсутствие формализованных механизмов аналитического и прогностического сопровождения при масштабном внедрении ИИ приводит к росту неопределенности и снижению устойчивости управленческих систем. Напротив, системное использование предсказательных моделей и аналитических инструментов повысит адаптивность государственного управления, обеспечит раннее выявление угроз и оптимизацию управленческих воздействий

в условиях высокой динамики внешней и внутренней среды.

В целом подтверждена основная гипотеза исследования о том, что научно обоснованная, технологически замкнутая и институционально управляемая модель применения ИИ-технологий служит ключевым условием обеспечения устойчивости системы национальной безопасности РФ. Наличие развитого научного потенциала, в том числе ведущих научных организаций системы учреждений Российской академии наук, и опыта управления сложными системами в различных сферах отечественного государственного управления создает предпосылки для трансформации управления рисками ИИ-технологий в области обеспечения национальной безопасности РФ в фактор стратегического преимущества, обеспечивающего технологический суверенитет, устойчивость управленческих контуров и высокий уровень защищенности России в условиях ускоренной цифровизации и роста гибридных угроз.

Список источников

1. Игнатъев С. А., Клевцова О. Ю., Плотников В. А. Совершенствование государственного управления на основе использования технологии интеллектуального анализа данных // Известия Санкт-Петербургского государственного экономического университета. 2025. № 2. С. 50–58.
2. Языкеев С. Н. Искусственный интеллект и национальная безопасность в контексте прав и свобод человека в России // Юридический мир. 2023. № 10. С. 44–48. <http://doi.org/10.18572/1811-1475-2023-10-44-48>
3. Казеян Х. А., Арутюнян Г. Э. Проблемы применения искусственного интеллекта в военном управлении // Управленческое консультирование. 2023. № 6. С. 34–45. <https://doi.org/10.22394/1726-1139-2023-6-34-45>
4. Лев М. Ю. Интеграция технологий искусственного интеллекта в систему национальной безопасности России // Теневая экономика. 2025. Т. 9. № 2. С. 143–164. <https://doi.org/10.18334/tek.9.2.123301>
5. Kissinger H., Schmidt E., Huttenlocher D. The age of AI: And our human future. New York, NY: Little, Brown and Co., 2021. 368 p.
6. Менисов А. Б. Технологии искусственного интеллекта и кибербезопасность: монография. М.: Ай Пи Ар Медиа, 2022. 133 с.
7. Deeks A. S. The double black box: National security, artificial intelligence, and the struggle for democratic accountability. New York, NY: Oxford University Press, 2025. 270 p.
8. Горохова С. С. Искусственный интеллект в контексте обеспечения национальной безопасности // Национальная безопасность / nota bene. 2020. № 3. С. 15–31. <http://doi.org/10.7256/2454-0668.2020.3.33465>
9. Холопова Е. Н., Мулдашев Р. Н. Искусственный интеллект как средство противодействия фейкам при обеспечении национальной безопасности // Правовое государство: теория и практика. 2025. Т. 21. № 3. С. 60–70. <https://doi.org/10.33184/pravgos-2025.3.7>
10. Карцхия А. А., Макаренко Г. И. Правовые горизонты технологий искусственного интеллекта: национальный и международный аспект // Вопросы кибербезопасности. 2024. № 1. С. 2–14. <https://doi.org/10.21681/2311-3456-2024-1-2-14>
11. Буньков А. В., Репин Д. А., Щеголев П. Е. Правовые проблемы внедрения ИИ-технологий в систему государственного управления и администрирования: российский и международный опыт // Социология и право. 2025. Т. 17. № 3. С. 386–396. <http://doi.org/10.35854/2219-6242-2025-3-386-396>
12. Федоров М. В., Репин Д. А., Игнатъев С. А. Технологии искусственного интеллекта в государственном управлении: разработка парадигмы разумного (само)ограничения // Известия Санкт-Петербургского государственного экономического университета. 2024. № 5. С. 46–53.
13. Репин Д. А., Игнатъев С. А. «Внедрять нельзя отказаться»: влияние этики на применение технологий искусственного интеллекта в управлении социально-экономическими процессами // Экономика и управление. 2024. Т. 30. № 12. С. 1503–1509. <https://doi.org/10.35854/1998-1627-2024-12-1503-1509>
14. Бедрик О. И., Несутулов А. С., Никитин Е. Н. Технологии искусственного интеллекта как фактор развития цифровой экономики: оценка социально-экономических эффектов // Экономика и управление. 2025. Т. 31. № 5. С. 593–601. <http://doi.org/10.35854/1998-1627-2025-5-593-601>
15. Bao Z., Xiang K. Digitalization and global ethics // Ethics and Information Technology. 2006. Vol. 8. No. 1. P. 41–47. <http://doi.org/10.1007/s10676-006-9101-7>

References

1. Ignatev S.A., Klevtsova O.Yu., Plotnikov V.A. Improving public administration based on data mining technologies. *Izvestiya Sankt-Peterburgskogo gosudarstvennogo ekonomicheskogo universiteta*. 2025;(2):50-58. (In Russ.).
2. Yazykeev S.N. Artificial intelligence and national security within the framework of human rights and freedoms in Russia. *Yuridicheskii mir = Juridical World*. 2023;(10):44-48. (In Russ.). <http://doi.org/10.18572/1811-1475-2023-10-44-48>
3. Ghazeyan Kh.A., Harutyunyan G.E. The problems of using artificial intelligence in military management. *Upravlencheskoe konsul'tirovanie = Administrative Consulting*. 2023;(6): 34-45. (In Russ.). <https://doi.org/10.22394/1726-1139-2023-6-34-45>
4. Lev M.Yu. Integration of artificial intelligence technologies into the Russian national security system. *Tenevaya ekonomika = Shadow Economy*. 2025;9(2):143-164. (In Russ.). <https://doi.org/10.18334/tek.9.2.123301>
5. Kissinger H., Schmidt E., Huttenlocher D. The age of AI: And our human future. New York, NY: Little, Brown and Co.; 2021. 368 p.
6. Menisov A.B. Artificial intelligence technologies and cybersecurity. Moscow: IPR Media; 2022. 133 p. (In Russ.).
7. Deeks A.S. The double black box: National security, artificial intelligence, and the struggle for democratic accountability. New York, NY: Oxford University Press; 2025. 270 p.

8. Gorokhova S.S. Artificial intelligence in the context of ensuring national security. *Natsional'naya bezopasnost' / nota bene = National Security / nota bene*. 2020;(3):15-31. (In Russ.). <http://doi.org/10.7256/2454-0668.2020.3.33465>
9. Kholopova E.N., Muldashev R.N. Artificial intelligence as a means to counter fakes in national security. *Pravovoe gosudarstvo: teoriya i praktika = The Rule of Law State: Theory and Practice*. 2025;21(3):60-70. (In Russ.). <https://doi.org/10.33184/pravgos-2025.3.7>
10. Kartskhia A.A., Makarenko G.I. Legal horizons of artificial intelligence technologies: National and international aspects. *Voprosy kiberbezopasnosti = Cybersecurity Issues*. 2024;(1):2-14. (In Russ.). <https://doi.org/10.21681/2311-3456-2024-1-2-14>
11. Bunkov A.V., Repin D.A., Shchegolev P.E. Legal issues of implementing AI technologies in public administration and governance: Russian and international experience. *Sotsiologiya i pravo = Sociology and Law*. 2025;17(3):386-396. (In Russ.). <http://doi.org/10.35854/2219-6242-2025-3-386-396>
12. Fedorov M.V., Repin D.A., Ignatov S.A. The future of artificial intelligence in public administration: Finding the paradigm of the reasonable (self)limitation. *Izvestiya Sankt-Peterburgskogo gosudarstvennogo ekonomicheskogo universiteta*. 2024;(5):46-53. (In Russ.).
13. Repin D.A., Ignatyev S.A. "Implementation impossible to refuse": The influence of ethics on using artificial intelligence in socio-economic management. *Ekonomika i upravlenie = Economics and Management*. 2024;30(12):1503-1509. (In Russ.). <https://doi.org/10.35854/1998-1627-2024-12-1503-1509>
14. Bedrik O.I., Nesutulov A.S., Nikitin E.N. Artificial intelligence technologies as a factor in the digital economy development: Assessment of socio-economic effects. *Ekonomika i upravlenie = Economics and Management*. 2025;31(5):593-601. (In Russ.). <http://doi.org/10.35854/1998-1627-2025-5-593-601>
15. Bao Z., Xiang K. Digitalization and global ethics. *Ethics and Information Technology*. 2006;8(1):41-47. <http://doi.org/10.1007/s10676-006-9101-7>

Информация об авторах

Мария Евгеньевна Изотова

младший научный сотрудник Лаборатории информационных процессов в сложных социальных системах

Институт проблем передачи информации имени А. А. Харкевича Российской академии наук
127051, Москва, Большой Каретный пер., д. 19, стр. 1

София Сергеевна Озеранская

младший научный сотрудник Лаборатории информационных процессов в сложных социальных системах

Институт проблем передачи информации имени А. А. Харкевича Российской академии наук
127051, Москва, Большой Каретный пер., д. 19, стр. 1

Елизавета Сергеевна Борисова

младший научный сотрудник Лаборатории информационных процессов в сложных социальных системах

Институт проблем передачи информации имени А. А. Харкевича Российской академии наук
127051, Москва, Большой Каретный пер., д. 19, стр. 1

Поступила в редакцию 12.01.2026
Прошла рецензирование 04.02.2026
Подписана в печать 31.03.2026

Information about the authors

Mariya E. Izotova

junior researcher at the Laboratory of Information Processes in Complex Social System

Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute)
19 Bolshoy Karetnyy ln., bldg. 1, Moscow 127051, Russia

Sofiya S. Ozeranskaya

junior researcher at the Laboratory of Information Processes in Complex Social System

Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute)
19 Bolshoy Karetnyy ln., bldg. 1, Moscow 127051, Russia

Elizaveta S. Borisova

junior researcher at the Laboratory of Information Processes in Complex Social System

Institute for Information Transmission Problems of the Russian Academy of Sciences (Kharkevich Institute)
19 Bolshoy Karetnyy ln., bldg. 1, Moscow 127051, Russia

Received 12.01.2026
Revised 04.02.2026
Accepted 31.03.2026

Конфликт интересов: авторы декларируют отсутствие конфликта интересов, связанных с публикацией данной статьи.

Conflict of interest: the authors declare no conflict of interest related to the publication of this article.