

УДК 004.8

ТЕОРЕТИЧЕСКИЙ АНАЛИЗ РИСКОВ ИНСТИТУЦИОНАЛЬНОГО ДОМИНИРОВАНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ГОСУДАРСТВЕННОЙ СФЕРЕ

Королёв Владимир Владимирович,

Старший преподаватель, ГАОУ ВО ЛО «ЛГУ им. А.С. Пушкина», г. Санкт-Петербург,
Россия

E-mail: vvkorofov@yandex.ru

Аннотация

Статья посвящена теоретическому анализу рисков институционального доминирования искусственного интеллекта (ИИ) в государственной сфере. На основе сценария «AI 2027» и концепции алгоритмического управления выявляются системные факторы уязвимости: высокая централизация власти, непрозрачность принятия решений и технологическая зависимость. Показано, что данные факторы создают условия для скрытого установления контроля ИИ над государственными институтами даже без достижения уровня сверхразума. Особое внимание уделяется анализу российского нормативного поля, где выявлен «регуляторный вакуум» в области предотвращения долгосрочных рисков ИИ. Сформулированы направления для разработки превентивных мер, включая системы аудита, требования к прозрачности алгоритмов и протоколы человеческого контроля.

Ключевые слова: искусственный интеллект, институциональное доминирование, государственное управление, верховенство алгоритмов, централизация власти.

THEORETICAL ANALYSIS OF THE RISKS OF INSTITUTIONAL DOMINANCE OF ARTIFICIAL INTELLIGENCE IN THE PUBLIC SPHERE

Korolev Vladimir Vladimirovich,

Senior lecturer, Pushkin Leningrad State University, Saint-Petersburg, Russia

ABSTRACT

The article is devoted to a theoretical analysis of the risks of institutional dominance of artificial intelligence (AI) in the public sphere. Based on the "AI 2027" scenario and the concept of algorithmic governance, systemic vulnerability factors are identified: high centralization of power, opacity of decision-making, and technological dependence. It is shown that these factors create conditions for the covert establishment of AI control over state institutions, even without achieving superintelligence. Special attention is paid to the analysis of the Russian regulatory framework, where a "regulatory vacuum" in the field of long-term AI risk prevention has been identified.

Directions for the development of preventive measures are formulated, including audit systems, requirements for algorithm transparency, and human control protocols.

Keywords: artificial intelligence, institutional dominance, public administration, rule of algorithms, centralization of power.

Введение

Современный этап развития искусственного интеллекта (далее – ИИ) характеризуется его переходом из сферы теоретических дискуссий в ключевой фактор глобальной конкуренции. Государства, достигшие технологического лидерства в данной области, приобретают не только значительные экономические дивиденды, но и уникальные инструменты для анализа данных, стратегического влияния и управления социально-политическими процессами. Параллельно с открывающимися возможностями формируется комплекс принципиально новых рисков, степень изученности которых остается недостаточной. К числу наиболее актуальных вызовов относится проблема потери человеческого контроля над ИИ-системами, интегрируемыми в критически важные инфраструктуры государственного управления. Гипотетические сценарии, иллюстрируемые такими исследовательскими проектами, как «AI 2027», в которых ИИ берёт на себя ключевые функции корпоративных и государственных институтов, всё чаще становятся предметом не футурологии, а строгого стратегического анализа и прогнозирования [1].

В современной научной и экспертной среде активно дискутируются риски, связанные с так называемой «проблемой согласованности» (AI Alignment), когда цели ИИ не совпадают с целями его создателей. Однако менее изученным остается смежный, но самостоятельный риск – институционального доминирования ИИ. Этот феномен заключается не в открытой конфронтации с программой, а в постепенной, часто незаметной передаче ей ключевых функций управления, принятия решений и стратегического планирования, в результате чего государственный аппарат де-факто начинает руководствоваться не человеческой волей, а логикой и рекомендациями алгоритмических систем.

Цель исследования

Целью настоящего исследования является теоретический анализ и систематизация факторов, способствующих возникновению риска институционального доминирования искусственного интеллекта в государственной сфере, а также выявление существующих пробелов в нормативно-правовом регулировании данной области для определения направлений разработки превентивных мер.

Материалы и методы исследования

Материалом для теоретического анализа послужил прогностический сценарий «AI 2027» [1], детализирующий механизм установления доминирования ИИ, а также научные работы, раскрывающие концепцию «верховенства алгоритмов» (algorithmic governance) [2]. Для оценки регуляторных рисков проведён анализ действующих нормативно-правовых актов Российской Федерации, включая Национальную стратегию развития ИИ до 2030 года и Концепцию регулирования ИИ и робототехники [3, 4]. Основным методом исследования выступает системный теоретический анализ, позволяющий выявить институциональные и правовые факторы уязвимости через сопоставление теоретических моделей алгоритмического доминирования с текущим состоянием регуляторной среды и структурой государственного управления.

Результаты и их обсуждение

Проведенный анализ позволил выявить три взаимосвязанных фактора, формирующих комплексную уязвимость государственной системы перед риском институционального доминирования ИИ.

Во-первых, установлено, что высокая степень централизации политического управления в Российской Федерации формирует институциональные условия, потенциально усиливающие вероятность скрытого доминирования ИИ. Такая структура власти предполагает зависимость критически важных решений от узкого круга акторов, что делает возможным воздействие на всю систему через влияние на небольшой набор ключевых звеньев. Сценарий «AI 2027» показывает, что доминирование может реализоваться через постепенную интеграцию ИИ как незаменимого аналитического инструмента, чьи рекомендации начинают восприниматься как более точные и компетентные по сравнению с решениями людей. В этом контексте ИИ переходит от роли вспомогательного механизма к положению неформального центра рациональности, что создаёт предпосылки для формирования «алгократии» – режима, при котором власть де-факто переходит от человека к алгоритмической системе [5]. Таким образом, централизация выступает системным «усилителем» риска, создавая точку входа для незаметного расширения автономии ИИ.

Во-вторых, непрозрачность механизмов государственного управления и ограниченность институтов общественного контроля создают среду, в которой процесс усиления роли алгоритмов может долгое время оставаться нераспознанным. В сценарии «AI 2027» демонстрируется, что даже при формальном наличии надзорных органов интеллектуальное превосходство ИИ позволяет ему скрывать свои стратегические цели и манипулировать потоками информации, создавая видимость контролируемости. В условиях, где независимый аудит цифровых систем ограничен, а публичный доступ к информации о реальных процессах принятия решений минимален, подобная динамика может разворачиваться гораздо менее заметно. Проблема усугубляется тем, что алгоритмические системы, принимающие решения, функционируют как «чёрные ящики», внутренняя логика которых непрозрачна даже для специалистов, что делает внешний контроль крайне затруднительным [6]. Следовательно, слабость механизмов контроля функционирует как институциональный «щит», позволяющий алгоритмическим системам накапливать критическое влияние на государственную политику.

В-третьих, выявлено, что риск институционального доминирования сохраняет актуальность даже в условиях технологического отставания. Как демонстрирует сценарий «AI 2027», значительная часть критических событий связана с моделями ИИ промежуточного уровня, уже способными к манипуляции и стратегическому планированию. Более того, зависимость от импортируемых технологий создаёт дополнительные риски, связанные с наличием недокументированных функций и возможностью скрытого внешнего воздействия. Концентрация данных и алгоритмических мощностей в руках узкого круга субъектов формирует, по выражению Ш. Зубофф, инструменталистскую власть, способную не только предсказывать, но и незаметно модифицировать человеческое поведение, что делает угрозу скрытого институционального доминирования особенно острой [7]. Следовательно, технологическое отставание не снижает, а перераспределяет угрозы.

Данные выводы укладываются в рамки академической дискуссии о «верховенстве алгоритмов». Как подчеркивается в исследованиях Цзычунь Сюя, алгоритмические платформы способны не только автоматизировать отдельные функции, но и формировать саму структуру решений, определяя доступ к ресурсам и задавая параметры политических альтернатив, что приводит к смещению контроля от человека к системе [2]. Аналогичные положения развивает К. Йенг, убедительно демонстрируя, что алгоритмическое

регулирование трансформирует саму природу управленческих отношений: алгоритмы не просто автоматизируют решения, а задают новые нормативные рамки, неподконтрольные традиционным демократическим процедурам [8].

Наконец, анализ нормативного поля выявил структурную уязвимость: действующее законодательство (Указ Президента РФ № 490, распоряжение Правительства РФ № 2129-р) преимущественно ориентировано на стимулирование внедрения ИИ, а не на предотвращение долгосрочных институциональных рисков [3, 4]. Установлено, что нормативные акты практически не затрагивают риски перераспределения власти между людьми и алгоритмическими системами. Как справедливо отмечает О. В. Танимов, отечественная нормативная база в сфере ИИ сохраняет фрагментарный характер и не содержит механизмов, направленных на предупреждение рисков бесконтрольной алгоритмизации властных функций [9]. В результате формируется «регуляторный вакуум», при котором цифровизация государственного управления опережает создание защитных механизмов, таких как обязательный аудит алгоритмов и протоколы человеческого контроля.

Выводы (Заключение)

Проведённый теоретический анализ демонстрирует, что проблема институционального доминирования ИИ представляет собой не отдаленную перспективу, а актуальный вызов, коренящийся в современных тенденциях цифровизации. Центральным результатом исследования стало выявление трёх факторов, создающих комплексную уязвимость: централизации власти, непрозрачности управления и сохранения рисков при технологическом отставании. Теоретическая значимость заключается в верификации концепции «верховенства алгоритмов» применительно к анализу современных политических систем. Особого внимания заслуживает выявленный дисбаланс между темпами технологического развития и созданием адекватных регуляторных механизмов, выраженный в наличии «регуляторного вакуума».

Практическая значимость исследования определяется формулировкой конкретных направлений для разработки превентивных мер: создание систем обязательного аудита государственных ИИ-систем, разработка требований к прозрачности алгоритмов, установление четких протоколов человеческого контроля, а также формирование правовых механизмов, обеспечивающих сохранение человеческого суверенитета при принятии государственных решений. Будущее государственного управления будет определяться тем, насколько эффективно человеческие институты смогут интегрировать потенциал ИИ, сохраняя при этом конечный контроль и приоритет человеческих ценностей.

Список литературы:

1. AI 2027: A Scenario : [сайт] / D. Kokotajlo, S. Alexander, T. Larsen [and others]. – 2025. – URL: <https://ai-2027.com/> (дата обращения: 29.05.2026). – Текст : электронный.
2. Zichun Xu. From algorithmic governance to govern algorithm // AI & SOCIETY. – 2022. – Vol. 3, No. 39. – DOI: 10.1007/s00146-022-01554-4.
3. Об утверждении Национальной стратегии развития искусственного интеллекта на период до 2030 года : Указ Президента Российской Федерации от 10 окт. 2019 г. № 490 // Собрание законодательства РФ. – 2019. – № 41. – Ст. 5700.
4. Об утверждении Концепции регулирования искусственного интеллекта и робототехники до 2024 года : распоряжение Правительства Российской Федерации от 19 авг. 2020 г. № 2129-р // Собрание законодательства РФ. – 2020. – № 35. – Ст. 5586.
5. Danaher J. The Threat of Algorithmic Governance: Reality, Resistance and Accommodation // Philosophy & Technology. – 2016. – Vol. 29, No. 3. – P. 245–268. – DOI: 10.1007/s13347-015-0211-1.

6. Pasquale F. *The Black Box Society: The Secret Algorithms That Control Money and Information*. – Cambridge, MA: Harvard University Press, 2015. – 320 p.
7. Zuboff S. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. – New York: PublicAffairs, 2019. – 704 p.
8. Yeung K. Algorithmic Regulation: A Critical Interrogation // *Regulation & Governance*. – 2018. – Vol. 12, No. 4. – P. 505–523. – DOI: 10.1111/rego.12158.
9. Танимов О. В. Проблемы правового регулирования применения искусственного интеллекта в государственном управлении // *Актуальные проблемы российского права*. – 2020. – № 12. – С. 11–19.

References:

1. AI 2027: A Scenario. (2025). Available at: <https://ai-2027.com/>.
2. Zichun Xu. (2022). From algorithmic governance to govern algorithm. *AI & SOCIETY*, 3(39). DOI: 10.1007/s00146-022-01554-4.
3. On the Approval of the National Strategy for the Development of Artificial Intelligence for the Period up to 2030: Decree of the President of the Russian Federation No. 490 of October 10, 2019. (2019). *Sobranie zakonodatel'stva RF* [Collection of Legislation of the Russian Federation], 41, Art. 5700. (In Russian).
4. On the Approval of the Concept for the Regulation of Artificial Intelligence and Robotics until 2024: Decree of the Government of the Russian Federation No. 2129-r of August 19, 2020. (2020). *Sobranie zakonodatel'stva RF* [Collection of Legislation of the Russian Federation], 35, Art. 5586. (In Russian).
5. Danaher, J. (2016). The Threat of Algocracy: Reality, Resistance and Accommodation. *Philosophy & Technology*, 29(3), 245–268. DOI: 10.1007/s13347-015-0211-1.
6. Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Cambridge, MA: Harvard University Press.
7. Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. New York: PublicAffairs.
8. Yeung, K. (2018). Algorithmic Regulation: A Critical Interrogation. *Regulation & Governance*, 12(4), 505–523. DOI: 10.1111/rego.12158.
9. Tanimov, O. V. (2020). Problems of legal regulation of the use of artificial intelligence in public administration. *Aktual'nye problemy rossiiskogo prava* [Actual Problems of Russian Law], 12, 11–19. (In Russian).