

Научная статья
УДК 34.09

Потенциал применения ИИ-агентов в контрольно-надзорной деятельности

Александр Сергеевич Киселев

Центр исследований и экспертиз Юридического факультета ФГБОУ ВО
«Финансовый университет при Правительстве Российской Федерации»,
Москва, Россия
alskiselev@fa.ru

Аннотация. Изучение потенциала применения искусственного интеллекта (ИИ) в контрольно-надзорной деятельности на сегодняшний день актуально из-за перспектив повышения эффективности работы государственных органов и необходимости регулирования использования технологий. При этом важно учитывать как перспективы, так и риски внедрения различных видов ИИ, а также разрабатывать нормативно-правовую базу. В отличие от классических ИИ-систем, которые обычно выполняют одну узкоспециализированную задачу (например, отвечают на вопросы), ИИ-агент действует проактивно, планирует свои шаги и взаимодействует с внешними сервисами. В контрольно-надзорной деятельности в перспективе целесообразно внедрение именно таких программ, тем не менее стоит обозначить ключевые риски и предложить механизмы правового регулирования.

Ключевые слова: право, контроль, надзор, информационные технологии, органы власти, контрольно-надзорная деятельность, искусственный интеллект, власть, государственное управление, государственный контроль

Original article

The potential of using AI agents in control and supervisory activities

Aleksandr S. Kiselev

Center for Research and Expertise at the Faculty of Law, Financial University under
the Government of the Russian Federation, Moscow, Russia
alskiselev@fa.ru

Abstract. Studying the potential of using artificial intelligence (AI) in control and supervisory activities is relevant today due to the prospects for improving the efficiency of government agencies and the need to regulate the use of technology. At the same time, it is important to take into account both the prospects and risks of implementing various types of AI, as well as to develop a regulatory framework. Unlike classical AI systems, which usually perform a single highly specialized task

(for example, answer questions), an AI agent acts proactively, plans its steps, and interacts with external services. In the future, it is advisable to introduce such programs in control and supervisory activities, however, it is worthwhile to identify the key risks and propose legal regulation mechanisms.

Keywords: law, control, supervision, information technology, authorities, control and supervisory activities, artificial intelligence, government, public administration, state control

Искусственный интеллект (ИИ) стремительно трансформирует цифровую экосистему любого современного государства, и центральное место в этой трансформации уже в ближайшем будущем займут ИИ-агенты. Под ИИ-агентом в современной научной парадигме понимается автономная или полуавтоматическая программа (или же система), способная воспринимать окружающую среду через данные, ставить цели и принимать решения для их достижения с минимальным вмешательством человека.

Их сущность заключается в сочетании трех ключевых свойств: автономности (способность действовать по заданным правилам без постоянного контроля со стороны человека-оператора), адаптивности (обучение на основе новых данных и опыта) и целенаправленности (выполнение конкретной задачи, от управления процессом до взаимодействия с пользователем).

Раскроем принципы работы ИИ-агента. В основе работы ИИ-агента лежит цикл восприятие – рассуждение – действие (рис. 1).

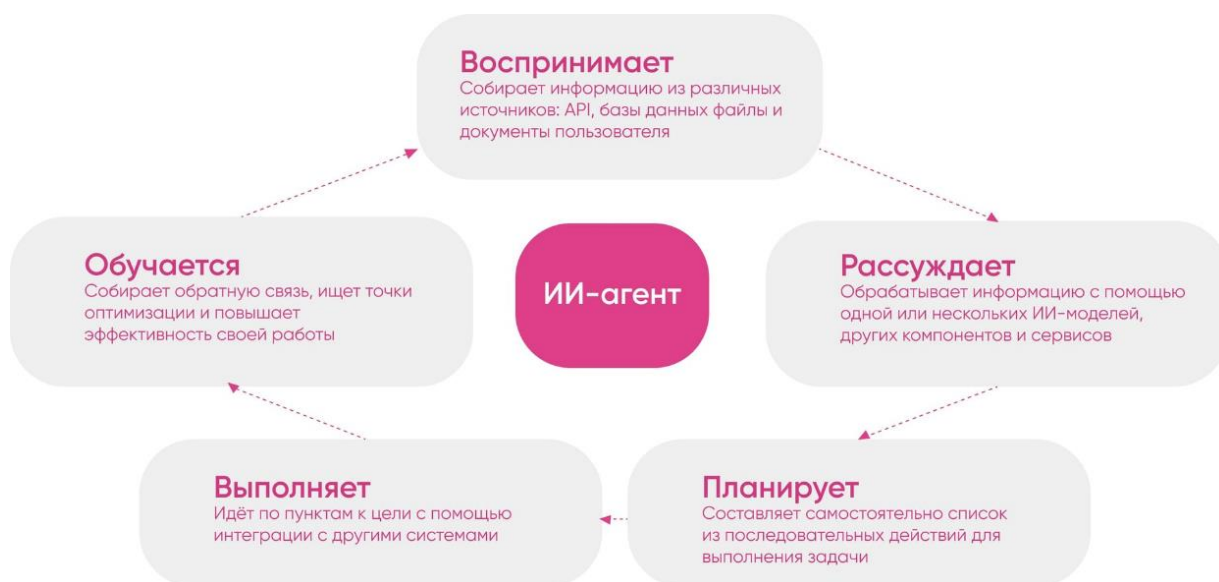


Рис. 1. Автономный цикл работы ИИ-агента [13]

Сначала агент получает информацию из внешней среды – звонок, документ или промт-запрос клиента. Потом он обрабатывает эти данные, анализирует их с помощью языковой модели и выстраивает цепочку решений. После этого выполняет нужные шаги: пишет письмо, проверяет звонок или

обновляет карточку в системе. Когда задача завершена, цикл повторяется – агент оценивает результат, делает выводы и постепенно становится точнее.

Внутри ИИ-агента работает несколько ключевых компонентов: память, профиль, планировщик, модуль рассуждения и модуль действий. Главными чертами ИИ-агента являются способность к самостоятельному планированию, наличие памяти для учёта контекста и возможность использовать внешние инструменты для взаимодействия с миром.

Эта функциональная мощь, однако, ставит самый сложный вопрос перед национальными правовыми системами мира: каков статус такого агента? Является ли он просто инструментом, расширением воли разработчика или оператора, или же его можно рассматривать как нового, цифрового, субъекта права?

В последнее десятилетие в науке доминировала мысль о том, что ИИ в принципе невозможно рассматривать в качестве субъекта права. Однако авторы все чаще говорят о наличии у современных ИИ-агентов черт, свойственных субъектам правоотношений. Например, Г. А. Маркосян отмечает, что «в контексте сделок особое значение приобретает способность ИИ-агентов не просто исполнять заранее запрограммированные алгоритмы, но и принимать решения в условиях неопределенности, руководствуясь заложенными в них целями и критериями оптимальности. Это качество сближает их с традиционными субъектами права» [7, с. 36].

Тем не менее подавляющее число исследователей, среди которых А. О. Мерцалова, придерживаются позиции о том, что «признание за искусственным интеллектом специальной правосубъектности, по всей видимости, должно быть отнесено к перспективе, когда технология достигнет высокой степени автономии и масштабного распространения, выходящего за рамки текущих ограниченных сценариев использования» [9, с. 912].

Сравнительный анализ подходов к правовому статусу ИИ-агентов

Инструментальный подход (преобладающая модель, включая Россию).

Большинство национальных правовых систем, включая Российскую Федерацию, придерживаются концепции, согласно которой ИИ-агент – это объект права (сложная компьютерная программа), а не субъект.

Соответственно ответственность за его действия полностью возлагается на физическое или юридическое лицо – разработчика, владельца или оператора. В некоторых случаях – правообладателя. В России этот подход закреплён в Стратегии развития искусственного интеллекта [2] и находит отражение в законопроектной работе. Например, «по итогам совещания с членами Правительства России от 5 апреля 2025 г. (Пр-706, пункт 5) правоохранительным органам было поручено проработать вопрос о целесообразности признания в качестве обстоятельства, отягчающего ответственность или наказание, использования при совершении правонарушения (преступления) технологий искусственного интеллекта и при необходимости обеспечить внесение в законодательство Российской Федерации соответствующих изменений» [1].

Приведём пример из практики, который наглядно свидетельствует о наличии инструментального подхода. Использование ИИ-агента «Альфа-Банком» (Россия) для одобрения микрозаймов через мобильное приложение. «AI-агент помогает управлять продуктами и финансами, решать нестандартные вопросы, фиксировать обращения и принимать благодарности. Он неотличим от человека: задаёт уточняющие вопросы, проявляет эмпатию, может подбодрить и пошутить. Клиенты 24/7 получают мгновенную поддержку по любым, даже сложным задачам. При этом существенно повышается скорость обработки запросов, снижается время ожидания ответа, что в свою очередь позитивно влияет на уровень удовлетворенности пользователей» [4]. В случае спорной ситуации (например, ошибочного отказа) ответственность несёт банк как оператор системы, а не алгоритм.

Аналогично в ЕС согласно Закону об искусственном интеллекте (AI Act, 2024) [16]. Системы ИИ ранжируются по степени риска, и для каждой группы вводится специальное регулирование – от самых строгих правил для высокорискованных систем до полной свободы для систем ИИ, не представляющих угрозу (рис. 2). Разработчики и поставщики систем высокого риска несут строгую ответственность за соответствие их продуктов требованиям безопасности и надёжности.

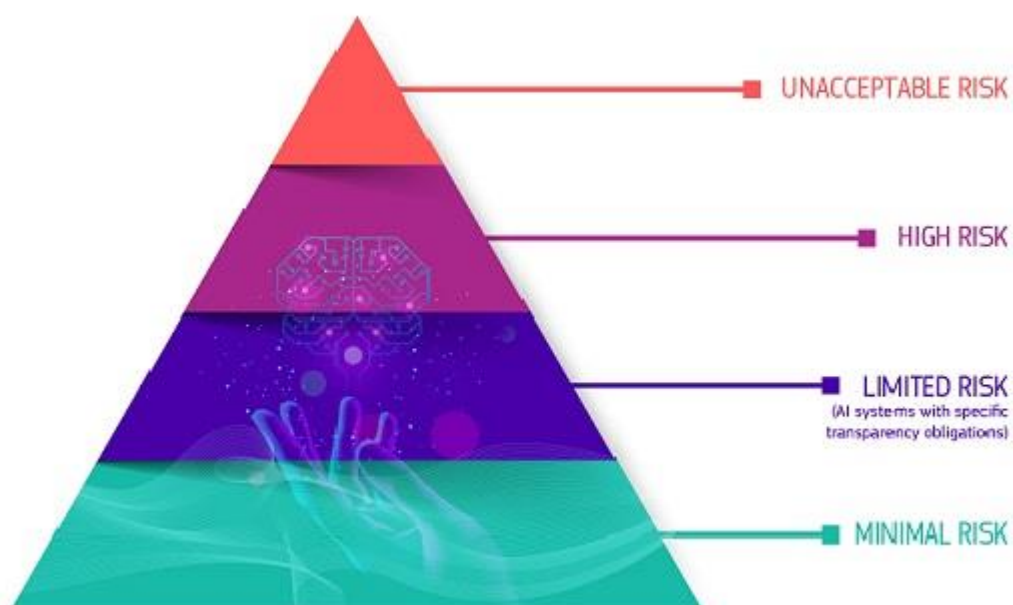


Рис. 2. Иерархия рисков, лежащая в основе Закона ЕС об искусственном интеллекте [15]

Экспериментальный подход (Саудовская Аравия, особые экономические зоны). Наиболее радикальный прецедент был создан в 2017 году, когда Саудовская Аравия предоставила гражданство человекоподобному роботу «София» от компании Hanson Robotics [10]. Хотя этот шаг носил в большей степени символический и рекламный характер, он спровоцировал глобальную дискуссию о возможности наделения продвинутых ИИ-агентов элементами правового статуса.

Более прагматичные инициативы имеются в нашей стране. Например, в СЭЗ «Алабуга» (Россия) тестируется правовой режим для беспилотного транспорта, где ИИ-агент, управляющий автомобилем, в рамках эксперимента признаётся ключевым «водителем», а ответственность распределяется между владельцем, оператором и производителем по сложной схеме, описанной в специальном регламенте. В этой связи в Республике Татарстан планируют выдавать сертификаты соответствия для искусственного интеллекта, применяемого на беспилотном транспорте.

А. В. Ларинская и А. К. Шереметьева убеждены, что «в будущем, вероятно, будут появляться специализированные институты, рассматривающие споры с участием высокоавтономных систем. Это могут быть как суды, так и квазисудебные органы, обладающие познаниями в сфере технологий. Наряду с этим продолжится развитие договорной модели, где ИИ выступает независимым сторонним агентом, обладающим электронной доверенностью. Тем самым право будет постепенно «приручать» новые технические феномены, трансформируя классические доктрины и дополняя их положениями, релевантными для цифровой эпохи» [6, с. 232].

Функциональный подход (США, Китай). В США и Китае правовое регулирование развивается по отраслевому принципу, фокусируясь не на абстрактном статусе ИИ-агента, а на конкретных последствиях его применения. Ключевые кейсы связаны с автономным оружием и киберфизическими системами. Например, Управление по санитарному надзору за качеством пищевых продуктов и медикаментов США (FDA) одобрило к применению ряд медицинских ИИ-агентов (в частности AI-Based Histologic Measurement of NASH (AIM-NASH)) для диагностики [14]. С правовой точки зрения такой агент рассматривается как медицинское изделие, подлежащее сертификации. Его «решения» – это результат работы алгоритма, а юридическую ответственность за ошибку диагностики несёт медицинское учреждение, использующее этот инструмент, и, потенциально, его производитель в рамках лицензионной ответственности.

На основе обобщения изученного материала мы можем сделать следующий вывод. Сущность ИИ-агента как автономного, целеориентированного инструмента пока не привела к признанию за ним самостоятельного правового статуса в большинстве стран мира. Доминирующим является инструментальный подход, где агент – объект права, а бремя ответственности лежит на конкретном человеке или организации, что прямо прописано в законе. Однако по мере усложнения агентов, особенно в гибридных киберфизических системах (беспилотники, автономные заводы, автономные системы управления техническими процессами), правовые системы будут вынуждены эволюционировать в соответствии с влиянием технического прогресса и ростом функционала ИИ-агентов. Вероятным сценарием представляется не предоставление агентам «прав» в человеческом понимании, а создание сложных, отраслевых правовых режимов, которые будут детально регулировать их разработку, внедрение и эксплуатацию, чётко

распределяя риски и ответственность между всеми участниками цифровой экосистемы.

Потенциала применения ИИ в контрольно-надзорной деятельности

Контрольно-надзорная деятельность, традиционно основанная на выборочных проверках и рутинном анализе данных, в настоящее время переживает трансформацию благодаря внедрению искусственного интеллекта (ИИ). ИИ-агенты открывают новые возможности для повышения эффективности, объективности и профилактической направленности надзора. Их применение позволяет перейти от наиболее распространённой модели «по факту нарушения» к проактивной, основанной на управлении рисками, т.е. на риск-ориентированном подходе к контролю и надзору.

Некоторые направления, где может применяться ИИ в контрольно-надзорной деятельности:

- 1) ИИ позволяет более точно определять потенциально опасные объекты, которые должны в первоочередном порядке подлежать проверке;
- 2) прогнозирование опасных ситуаций, например, компьютерное моделирование аварий, инцидентов и отказов оборудования, которое помогает направлять подконтрольному субъекту предписания о превентивном техническом обслуживании;
- 3) ИИ позволяет проводить одновременно несколько контрольно-надзорных мероприятий в отношении разных объектов. Например, при видеосъёмке компьютерная программа на основе ИИ может распознавать технические устройства и детали, объекты, помещения, сооружения, и выдавать аналитический отчёт о соблюдении или несоблюдении обязательных требований противопожарной и санитарной безопасности;
- 4) переход от реактивных к превентивным мерам (ИИ поможет фокусировать проверки на наиболее проблемных объектах, минимизируя давление на добросовестный бизнес).

Рассмотрим примеры из практики, где уже широко применяются ИИ-агенты.

В рамках налогового и финансового надзора осуществляется прогнозирование рисков уклонения (Россия). Федеральная налоговая служба (ФНС) России с 2017 года использует систему «АСК НДС-3», основанную на технологиях ИИ и машинного обучения. Агент анализирует цепочки сделок миллионов компаний в реальном времени, выявляя схемы незаконного возмещения НДС через фирмы-однодневки. «Система строит цифровые профили налогоплательщиков, оценивает риски и автоматически формирует подборки для проверок. По данным ФНС, это позволило увеличить выявляемость нарушений и сместить акцент на профилактику, предупреждая компании о рисках до начала проверки. Программный комплекс "АСК НДС-3" автоматически формирует отчет, который с ноября 2017 года является достаточным основанием для возбуждения уголовного дела в случае, если сумма нарушений превышает установленный законом размер» [5].

Экологический надзор (Китай, ЕС). В Китае внедряется национальная система экологического мониторинга, разработанная китайскими учёными

под руководством доктора Давэя Чжана из Национального центра экологического мониторинга Китая [11]. Система объединяет данные со спутников, дронов, наземных станций и морских датчиков. Она работает почти в реальном времени, собирая информацию о воздухе, воде, почве и океанах. Дроны и автоматические лаборатории собирают данные без участия человека, искусственный интеллект анализирует информацию и строит прогнозы, а центральная платформа объединяет все данные и выдаёт рекомендации.

ИИ-агенты интегрируют данные с тысяч датчиков контроля качества воздуха и воды, спутникового мониторинга и промышленных объектов. Алгоритмы в режиме 24/7 отслеживают превышение нормативов, прогнозируют распространение загрязнений и автоматически генерируют предписания для предприятий-нарушителей. Аналогичные, хотя и менее централизованные системы на основе ИИ для анализа спутниковых снимков и данных датчиков используются Европейским агентством по окружающей среде (англ. European Environment Agency, сокр. ЕЕА) для контроля за вырубкой лесов и промышленными выбросами.

Надзор за субъектами на финансовых рынках (США, Великобритания, Испания).

Комиссия по ценным бумагам и биржам США (SEC) применяет систему MIDAS (Market Information Data Analytics System)¹, а Управление по финансовому регулированию и надзору Великобритании (FCA) – инструменты на базе машинного обучения². ИИ-агенты анализируют миллионы сделок ежедневно, выявляя паттерны, характерные для инсайдерской торговли, манипулирования рынком (спуфинг³) или несанкционированной деятельности. Эти системы способны обнаруживать сложные, слабо заметные связи и аномалии, которые практически невозможно выявить ручными методами.

Помимо этого в Испании в 2018 году разработана компьютерная модель, основанная на искусственных нейронных сетях, которая определяет вероятность проявления коррупции в провинциях при осуществлении контрольной деятельности.

¹ Система была внедрена в январе 2013 года. SEC поручила построить MIDAS компании Tradeworx, которая разработала платформу для анализа рыночных данных. В 2020 году в статье Фреда Соммерса сообщалось, что данные MIDAS неполны. Например, в систему не включены транзакции, которые происходят не на бирже и не reported через три системы отчётности FINRA. Однако SEC заявляла, что MIDAS помогает улучшить понимание работы рынков и способствует эмпирическому пониманию гипотез о поведении рынка и регулировании рынка [18].

² По информации на ноябрь 2024 года, FCA активно работает с инновациями в области ИИ и, в частности, представило проект AI Lab. Вместе с Банком Англии регулятор изучает, как механизмы ИИ могут повлиять на финансовые рынки [17].

³ Термин из области сетевой безопасности, обозначающий кибератаку, при которой злоумышленник маскируется под другого человека, компанию или объект. Цель – завоевать доверие жертвы, чтобы получить доступ к устройствам, украсть деньги, личные данные или распространить вирусные программы.

«Модель позволяет прогнозировать коррупцию на период до трёх лет в зависимости от характеристик конкретного региона. Учёные учли такие факторы, как изменение налога на недвижимость, рост экономики, повышение цен на жильё, увеличение количества депозитных учреждений и нефинансовых фирм, слишком долгое нахождение у власти одной и той же партии и другие. Модель протестировали на 300 серьёзных делах, и результат оказался следующим: машина предсказала проявление коррупции с вероятностью 80 %. Однако учёные отмечали, что система не универсальна, так как обучена на испанском опыте, и её нельзя автоматически переносить на другие страны» [8].

В целом при позитивной картине следует помнить, что внедрение ИИ-агентов сопряжено с серьёзными вызовами:

- 1) риск алгоритмической предвзятости, недостаточная прозрачность принятия решений («чёрный ящик»), вопросы ответственности за автоматизированные решения. С целью нивелирования указанного риска С. В. Спрыгина предлагает «разрабатывать и внедрять механизмы аудита и проверки алгоритмов на предмет предвзятости, чтобы избежать таких негативных явлений» [12, с. 289];
- 2) зависимость от качества и репрезентативности обучающих данных, уязвимость к «состязательным атакам»;
- 3) необходимость переподготовки персонала и изменения управленческих процессов.

Заключение

Потенциал ИИ-агентов в контрольно-надзорной деятельности заключается в создании «интеллектуального» надзора, который является непрерывным, всеобъемлющим (только в отношении данных, подлежащих проверке согласно букве закона) и предиктивным. Реальные кейсы в налоговой, экологической и финансовой сферах демонстрируют значительный прирост в эффективности выявления нарушений и переход к риск-ориентированной модели.

Однако реализация этого потенциала требует параллельного развития нормативной базы, обеспечения алгоритмической справедливости и прозрачности, а также инвестиций в цифровую инфраструктуру и компетенции надзорных органов. Будущее надзора остается, на наш взгляд, за гибридной моделью, где стратегические решения остаются за человеком, а рутинный анализ, мониторинг и первичная оценка рисков делегируются ИИ-агентам. При этом важно, чтобы ИИ не подменял и не заменял деятельность должностного лица контрольного органа – окончательное решение о квалификации правонарушения принимает человек. Это правило, вполне очевидно, должно идти красной строкой через всё современное законодательство, посвященное правовому регулированию всей сферы ИИ.

Список источников

1. О проведении эксперимента по установлению специального регулирования в целях создания необходимых условий для разработки и внедрения технологий искусственного интеллекта в субъекте Российской Федерации – городе федерального значения Москве, об особенностях обработки персональных данных при формировании региональных составов данных и предоставления доступа к региональным составам данных и внесении изменений в статьи 6 и 10 Федерального закона «О персональных данных»: Федеральный закон от 24 апреля 2020 г. № 123-ФЗ: с изм. от 8 августа 2024 г. № 233-ФЗ // Рос. газ. – 2020. – № 92.
2. О развитии искусственного интеллекта в Российской Федерации: Указ Президента РФ от 10 октября 2019 г. № 490: с изм. от 15 февраля 2024 г. № 124 // Собр. законодательства Рос. Федерации. – 2019. – № 41, ст. 5700.
3. Законопроект № 885494-8 от 07.04.2025 г. «О внесении изменения в статью 63 Уголовного кодекса Российской Федерации» // Государственная Дума Федерального Собрания Российской Федерации: [сайт]. URL: <https://sozd.duma.gov.ru/bill/885494-8?ysclid=mlrwdq8fp3896865461> (дата обращения: 15.01.2026).
4. Альфа-Банк первым нанял на работу виртуального сотрудника для общения с клиентами // Альфа-банк: [сайт]. URL: <https://alfabank.ru/news/t/release/alfa-bank-pervim-nanyal-na-rabotu-virtualnogo-sotrudnika-dlya-obshcheniya-s-klientami/?ysclid=mlrwt985ur491442428> (дата обращения: 15.01.2026).
5. АСК НДС-3: как работает и кого проверяют первыми // Клерк. – 2018. – 18 января. URL: <https://www.klerk.ru/buh/articles/471297/> (дата обращения: 15.01.2026).
6. Ларинская А. В., Шереметьева А. К. Правовое регулирование и основные элементы правосубъектности юнитов искусственного интеллекта // Вопросы российского и международного права. – 2025. – Т. 15, № 1-1. – С. 221-236.
7. Маркосян Г. А. Правовой статус ИИ-агента: может ли нейросеть нести ответственность за сделку? // Юридический вестник Кубанского государственного университета. – 2025. – № 4. – С. 34-39. – DOI 10.31429/20785836-17-4-34-39.
8. Медведев Ю. Искусственный интеллект научился предсказывать коррупцию // РБК. – 2018. – 17 апреля. URL: https://rg.ru/2018/04/17/iskusstvennyj-intellekt-nauchilsia-predskazyvat-korrupciuu.html?utm_referrer=https%3A%2F%2Fyandex.ru%2F (дата обращения: 15.01.2026).
9. Мерцалова А. О. Искусственный интеллект как субъект и объект авторского права // Флагман науки. – 2025. – № 11 (34). – С. 911-914.
10. Миклашевская А. Саудовская Аравия предоставила гражданство роботу // Коммерсант. – 2017. – 29 октября. URL:

<https://www.kommersant.ru/doc/3454104?ysclid=mlrxahkntt600429568> (дата обращения: 15.01.2026).

11. Наговицын М. Китай создал самую мощную систему экологического мониторинга // Новости мира инноваций : [сайт]. URL: <https://innovanews.ru/info/news/ecology/kitajj-sozdal-samuju-moshhnuju-sistemu-ekologicheskogo-monitoringa/> (дата обращения: 15.01.2026).

12. Спрыгина С. В. Этика и регулирование искусственного интеллекта: баланс между инновациями и защитой прав человека // Трансформация бизнеса и общественных институтов в условиях цифровизации экономики : сборник научных трудов VII Национальной (российской) научно-практической конференции, Санкт-Петербург, 17–18 апреля 2025 года. – СПб. : Санкт-Петербургский университет технологий управления и экономики, 2025. – С. 287-290.

13. Шмалинюк К. ИИ-агенты: что это, их виды и какие задачи они решают // Марквиз. – 2025. – 26 июня. URL: <https://marquiz.ru/blog/ai-agents/> (дата обращения: 15.01.2026).

14. FDA впервые одобрила ИИ-инструмент для разработки лекарств от болезни печени // Ведомости. – 2025. – 10 декабря. URL: https://www.vedomosti.ru/health/medical_technology/news/2025/12/10/1162255-fda-odobrila-ii-instrument (дата обращения: 15.01.2026).

15. Risk hierarchy underpinning the EU AI Act // Департамент предпринимательства, торговли и занятости (DETE) Ирландии : [сайт]. URL: <https://enterprise.gov.ie/en/what-we-do/innovation-research-development/artificial-intelligence/eu-ai-act/> (дата обращения: 15.01.2026).

16. The European Union Artificial Intelligence Act, 12 July 2024 // The European Union : [сайт]. URL: <https://www.ey.com/content/dam/ey-unified-site/ey-com/en-gl/insights/public-policy/documents/ey-gl-eu-ai-act-07-2024.pdf?ysclid=mlrwzsk9os642455811> (дата обращения: 15.01.2026).

17. FCA становится более активной и инновационной в использовании ИИ. URL: <https://dzen.ru/a/ZyaFXdn6b2RPVJp8> (дата обращения: 15.01.2026).

18. Papers.ssrn.com : [сайт]. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3692464&__cf_chl_rt_tk=5aif90y2pPf1m6eOAaz45lgc90PRcJhsb.3QyDvmkqU-1771414480-1.0.1.1-M8D_e_QBvRVHKNy3bYkS3n5f49pOHot1fxSds5k8Jzc (дата обращения: 15.01.2026).

Статья поступила в редакцию 28.01.2026; одобрена 13.02.2026; принята к публикации 20.02.2026.

The article was submitted 28.01.2026; approved after reviewing 13.02.2026; accepted for publication 20.02.2026.