



Известия Саратовского университета. Новая серия. Серия: Экономика. Управление. Право. 2023. Т. 23, вып. 4. С. 428–438

Izvestiya of Saratov University. Economics. Management. Law, 2023, vol. 23, iss. 4, pp. 428–438

<https://eup.sgu.ru>

<https://doi.org/10.18500/1994-2540-2023-23-4-428-438>, EDN: ХСКАХР

Научная статья

УДК 338



Технологии искусственного интеллекта: классификация, ограничения, перспективы и угрозы

И. В. Вешнева

Саратовский государственный национальный исследовательский университет имени Н. Г. Чернышевского, Россия, 410005, г. Саратов, ул. Астраханская, д. 83

Вешнева Ирина Владимировна, доктор технических наук, профессор кафедры информационных систем и технологий в обучении, veshnevaiv@gmail.com, <https://orcid.org/0000-0002-3908-370X>

Аннотация. Введение. Решение задач повышения производительности труда, формирование прозрачности основных бизнес-процессов, создание новых производств требуют не только новых производственных технологий, но и организации информационных процессов, таких как сбор, хранение, обработка, анализ, выдачи ответа на запросы по информации, сопровождающей производственные процессы. Одним из наиболее перспективных инструментов решения этих задач являются технологии искусственного интеллекта. **Теоретический анализ.** Представлена классификация технологий искусственного интеллекта, выделены направления: машинное обучение, обработка естественного языка, компьютерное зрение, экспертные системы, усовершенствованное планирование, распознавание речи, робототехника. Дана характеристика основным технологиям внутри выделенной классификации. Для описания ограничений развития искусственного интеллекта использованы уровни ограничения развития физической реализации, безопасности использования, взаимодействия с окружением, уровень признания, возможности самоактуализации. Перспективы и риски структурированы как наборы аналогичных уровней, дополненные уровнем анализа в области энергетики. **Заключение.** Выявлено, что для обеспечения развития Общества 5.0 необходимо создание инновационных площадок и мегарегиональных кластеров для кооперации органов власти, предпринимателей, исследовательских центров в области технологий искусственного интеллекта.

Ключевые слова: искусственный интеллект, киберфизические системы, риски, безопасность, машинное обучение, обработка естественного языка, компьютерное зрение, экспертные системы, усовершенствованное планирование, распознавание речи, робототехника, устойчивое развитие России, Общество 5.0

Благодарности: Исследование выполнено при финансовой поддержке Российского фонда фундаментальных исследований (проект № 22-010-00465).

Для цитирования: Вешнева И. В. Технологии искусственного интеллекта: классификация, ограничения, перспективы и угрозы // Известия Саратовского университета. Новая серия. Серия: Экономика. Управление. Право. 2023. Т. 23, вып. 4. С. 428–438. <https://doi.org/10.18500/1994-2540-2023-23-4-428-438>, EDN: ХСКАХР

Статья опубликована на условиях лицензии Creative Commons Attribution 4.0 International (CC-BY 4.0)

Article

Artificial intelligence technologies: Classification, limitations, prospects and threats

I. V. Veshneva

Saratov State University, 83 Astrakhanskaya St., Saratov 410012, Russia

Irina V. Veshneva, veshnevaiv@gmail.com, <https://orcid.org/0000-0002-3908-370X>

Abstract. Introduction. Solving the problems of increasing labor productivity, creating transparency of key business processes, and creating new production facilities require not only new production technologies, but the organization of information processes, such as collection, storage, processing, analysis, and issuing responses to requests for information accompanying production processes. One of the most promising tools for solving these problems is AI technologies. **Theoretical analysis.** A classification of artificial intelligence technologies is presented, the following areas are highlighted: machine learning, natural language processing, computer vision, expert systems, advanced planning, speech recognition, robotics. The characteristics of the main technologies within the selected classification are given. To describe the limitations of the artificial intelligence development, the following levels of development limitation were used: physical implementation, safety of use, interaction with the environment, level of recognition, and the possibility of self-actualization. Prospects and risks are structured as sets of similar levels, complemented by a level of energy analysis. **Conclusion.** It has been revealed that to ensure the development of Society 5.0, it is necessary to create innovative platforms and mega-regional clusters for cooperation between authorities, entrepreneurs, and research centers in the field of artificial intelligence technologies.

Keywords: artificial intelligence, cyber-physical systems, risks, security, machine learning, natural language processing, computer vision, expert systems, advanced planning, speech recognition, robotics, sustainable development in Russia, Society 5.0

Acknowledgements: This work was supported by the Russian Foundation for Basic Research (project No. 22-010-00465).



For citation: Veshneva I. V. Artificial intelligence technologies: Classification, limitations, prospects and threats. *Izvestiya of Saratov University. Economics. Management. Law*, 2023, vol. 23, iss. 4, pp. 428–438 (in Russian). <https://doi.org/10.18500/1994-2540-2023-23-4-428-438>, EDN: XCKAXR
This is an open access article distributed under the terms of Creative Commons Attribution 4.0 International License (CC-BY 4.0)

Введение

В рамках Национальной программы «Цифровая экономика Российской Федерации» действует федеральный проект «Искусственный интеллект», ставящий своей основной задачей создание условий для широкого внедрения продукции, базирующейся на отечественных технологиях искусственного интеллекта (ИИ) [1]. Однако сегодня уровень внедрения ИИ в России в три раза ниже, чем в мире [2]. Вместе с тем в ключевых отраслях промышленности более 20% российских компаний внедрили технологии ИИ, что принесло рост рентабельности на 5% и позволит увеличить рост ВВП России на 1% к 2025 г. Однако более 50% опрошенных собственников компаний не знают о технологиях ИИ [2]. Решение задач повышения производительности труда, формирование прозрачности основных бизнес-процессов, создание новых производств требуют не только новых производственных технологий, но и организации информационных процессов, таких как сбор, хранение, обработка, анализ, выдачи ответа на запросы по информации, сопровождающей производственные процессы. Одним из наиболее перспективных инструментов решения этих задач являются технологии ИИ. Их применение демонстрирует значительный экономический эффект в здравоохранении [3], торговле, транспорте, логистике, автоматизированном производстве, банковском деле и т. д. [4]. Многие страны разрабатывают или внедряют собственные стратегии использования и развития ИИ [5]. В численности научных работ, посвященных проблемам разработки систем ИИ, наблюдается 8-кратный рост за последние пять лет. В работе [6], посвященной обзору технологий искусственного интеллекта и машинного обучения, проводится анализ препятствий на пути внедрения ИИ в экономику, в том числе ресурсных.

Цель исследования – провести классификацию развитых технологий искусственного интеллекта, выявить и структурировать возможные ограничения роста технологий ИИ, перспективы внедрения в различные сферы социально-экономических систем и возможные риски, особое внимание уделив специфике развития России.

Для реализации поставленной цели предложен оригинальный подход к разработке струк-

турирования оценок показателей внедрения технологий ИИ. Основываясь на предположении, что ИИ имитирует мышление человека, отдельные показатели применения ИИ можно систематизировать на основе трехгранной пирамиды: ограничения, перспективы, угрозы. Для понимания сложности круга анализируемых проблем в статье представлена классификация технологий ИИ.

Теоретический анализ

Классификация технологий искусственного интеллекта. Определимся с основной терминологией. ИИ понимается как свойство искусственных интеллектуальных систем выполнять творческие функции, традиционно считающиеся прерогативой человека [7]. В структуре ИИ можно выделить следующие направления.

- Машинное обучение [8].
- Компьютерное зрение [9].
- Обработка естественного языка [10].
- Экспертные системы [10, 11].
- Усовершенствованное планирование.
- Распознавание речи.
- Робототехника [8, 12].

Применение различных технических программных средств во многих случаях может быть подвергнуто классификации весьма условно. Это позволяет утверждать, что возможности развития ИИ следует рассматривать с различных аспектов. Одним из значимых аспектов классификации может быть область использования ИИ [13]. Отмечено, что методы ИИ наиболее часто используются в медицине, сельском хозяйстве и образовании [6, 13]. Однако требуется ускорить использование методов ИИ в социально-экономических системах [1, 2].

Основываясь на предположении, что ИИ имитирует мышление человека, отдельные показатели применения ИИ можно систематизировать на основе трехгранной пирамиды: ограничения, перспективы, угрозы. Каждая грань структурирована по принципу, подобному структуре пирамиды Маслоу, исключая эстетические потребности. В таблице представлены структура и аналогии с известной пирамидой Маслоу. Нумерация проведена снизу вверх для подчеркивания обосновывающей аналогии.



**Структура ограничений, перспектив и угроз внедрения ИИ
в различные сферы социально-экономических систем**

*Table. Structure of limitations, prospects, and threats of introducing AI into various areas
of socio-economic systems*

№	Пирамида Маслоу: потребности	Ограничения систем ИИ	Перспективы систем ИИ	Угрозы развития и использования систем ИИ
6	Самоактуализации	Функционируют в собственной сконструированной реальности	Создание специфического интеллектуального пространства, взаимодействующего с метавселенной	Трансформация языков и мышления, поиск экзистенциальных и онтологических смыслов
5	Познавательные	Перепрограммирование, «короткая» память систем ИИ	Достижение успеха и высокой оценки, признания	Требуется гарантия положительного успеха внедрения ИИ
4	В уважении	Уровень признания, достижения успеха, служебный рост	Системы коммуникаций цифровых двойников типа объект-объектного взаимодействия	Ядра ИИ могут организовать свой цифровой рынок с собственными правилами конкурентной борьбы
3	В принадлежности к группе	Уровень взаимодействия с окружением – модели взаимодействия киберфизических систем	Безопасность, реализуемая через киберфизические системы и коботы	Риски противоправных операций с личными данными, возможность манипулирования системами жизнеобеспечения
2	В безопасности	Безопасность данных	Промышленность как ключевой драйвер развития ИИ	Разрыв промышленных связей
1	Физиологические	Физическая реализация	Энергетика: качественно новые характеристики надежности, эффективности, доступности и управляемости	Экологические: существенный углеродный след систем ИИ

Рассмотрим данную структуру подробнее.

Ограничения систем искусственного интеллекта

В середине XX в. первые исследователи технологий ИИ выражали уверенность, что через 10 лет появятся интеллектуальные машины, способные мыслить на уровне человека. Однако внутренние ограничения развития технологий ИИ привели к замораживанию в развитии отрасли. Одним из многообещающих проектов последнего десятилетия выступила идея беспилотных транспортных средств. Например, Business Insider в 2016 г. предсказывал, через три–четыре года, уже к 2020-му по дорогам будут мчаться 10 млн беспилотных автономных такси. Эти предсказания пока не сбылись. Завышенные ожидания быстрых высоких результатов в отношении прогресса технологий ИИ опасно. Если развитие технологий ИИ не оправдает серьезных инвестиций в прогнозируемые короткие сроки, то может наступить новый отток средств и исследований из отрасли, что приведет к новому периоду «зимы ИИ». Попробуем выделить основные направления, в которых бурно развивающиеся технологии ИИ испытывают ограничения. Для создания

предварительной классификации ограничений используем структуру пирамиды Маслоу.

1. *Уровень физической реализации.* Физические, внутренние ограничения связаны с большим разрывом между возможностями ИИ и человеческого мозга. Человеческий мозг средним объемом 1500 см³ содержит около 90 млрд нейронов, это около 17 млрд узлов сети в 1 см³, что значительно превышает мощность существующий искусственной нейронной сети (ИНС). Самые мощные из действующих сетей не превышают 1 мм³ мозга человека по количеству узлов. Кроме того, важно понимать, что ИИ – это только модель мозга, которая работает не так, как мозг. Сложность модели не позволяет понять ход процесса принятия решений системой. Обучение нейронных сетей требует больших вычислительных и энергетических ресурсов. Исследователи [14] утверждают, что сегодня развитие традиционного ИИ подошло к пределу своих концептуальных возможностей, поэтому требуется разработка альтернативных алгоритмов и подходов к созданию интеллектуальных систем, например, квантовые вычисления, распределенные системы, нейроморфные системы



2. *Безопасность использования ИИ связана с данными.* Для того чтобы система ИИ смогла принимать решения, нужно спроектировать набор возможных решений, предоставить системе данные и «научить» принимать правильные решения. Собрать данные, структурировать и предоставить их в объеме, точно соответствующем требованиям решаемой задачи, – это трудоемкий процесс. Данные редко бывают «хорошими». Если систему обучить на таких данных, можно получить не вполне ожидаемые результаты. Например, компания Twitter создала чатбота Тэй с аватаркой и манерами 19-летней американки. За первые сутки общения чатбота с пользователями сети системы ИИ бот Тэй превратилась в расиста и сексиста. Созданный специалистами фирмы Hanson Robotics в 2016 г. человекоподобный робот София в публичном интервью со своим создателем неожиданно заявила о намерении «уничтожить человечество» [15]. Существует опасение, что ИИ не способен делать четкие, позитивно интерпретируемые выводы об объективной реальности, он создает свою собственную сконструированную реальность, которая служит основой для ограниченных выводов. Кроме того, обучение систем ИИ всегда строится на ранее собранных базах данных, экстраполяция которых приводит к ошибкам, которые невозможно отследить в ИНС.

Кроме качества и количества поступающих данных, можно выделить следующие ограничения: хрупкость (система неспособна адаптироваться к изменениям входных данных), забывание (при изучении ИНС новой задачи старые задачи забываются), сконструированная реальность (ИНС не способны оценивать обстоятельства и действовать в соответствии с ними), рискогенность (неспособность ИНС успешно действовать в ситуациях с высокой степенью риска). Также обнаружено, что системы ИИ легко обмануть. Например, изменение цветов фрагментов изображения может привести к полному непониманию системой ИИ изображения. Для программы исследуемые объекты не являются объектами и явлениями, они являются только программным кодом. Для решения различных проблем, связанных с непониманием системами ИИ объективной реальности, предложено использовать цифровых двойников киберфизических систем (см., например, [16]).

Киберфизические системы являются новыми технологиями (*enabling technologies*), позволяющими соединить искусственный, виртуальный, созданный техническими средствами и реальный физический мир, в котором возможно

реализовать взаимодействие распределенных систем технических объектов друг с другом физическими и программными методами. Переход от встроенных в различные объекты управляющих систем к киберфизическим системам, который в документах «проекта будущего» называют четвертой промышленной революцией, приведет как к взаимодействию ИИ – реальный мир, так и к прямому взаимодействию объект ИИ – объект ИИ [17].

3. *Уровень взаимодействия с окружением.* В социально-экономической структуре требуется создать модели взаимодействия киберфизических систем. Будем полагать, социальные коммуникации «субъект-субъект» выстраиваются на моделировании восприятия одним субъектом действий другого субъекта, субъекта или эмпатии. Отношения «субъект-объект» более 50 лет активно исследуются в работах по человеко-машинному взаимодействию, эргатическим системам, киберфизическим системам, которые основаны на моделировании сценариев взаимодействия, выявлении продукционных правил действие – следствие [18]. Отношения «объект-объект» в среде взаимодействующих систем ИИ требует сложного моделирования, понимания свойств объектов – технических систем ИИ. При этом сложность модели каждой из взаимодействующих систем не позволяет понять ход процесса принятия решений отдельной системой. Не существует развернутых исследований, алгоритмов, шаблонов, описывающих взаимодействие систем ИИ. Важно отметить, что разработка алгоритмов взаимодействия человека и технических систем проводится несколько десятков лет и является молодым, активно развивающимся направлением исследований, в то время как взаимодействие систем ИИ исследовалось только на уровне технических и программных интерфейсов. Каждая система ИИ создает свою собственную реальность, которая служит основой для выводов. Развитие взаимодействия «объект-объект» будет связано с совмещением различных ограниченных проекций реальности и может приводить к конфликтности искусственных реальностей. Процессы обучения технических систем техническими системами имеют слабо предсказуемые последствия, могут возникнуть и происходить стремительно.

4. *Уровень признания, достижения успеха, служебный рост.* Признание результатов работы систем ИИ также является ограничением их использования. Например, экспериментальное внедрение ИИ Deep Patient в американских медицинских клиниках на историях болезни 700 тыс.



пациентов позволило натренировать ИИ прогнозировать заболевания, например онкологию. Параллельно система научилась успешно решать труднейшую для медиков задачу диагностики такого психического заболевания, как шизофрения. При этом ни медицинский персонал, ни создатели системы ИИ так и не смогли понять, по каким признакам Deep Patient абсолютно точно вычислял будущих шизофреников. Законность лечения на основе диагноза, который получен по неизвестным признакам и алгоритмам, не установлена. Легитимность и обоснованность решений на основе результатов, полученных методами ИИ, являются проблемой во всех сферах деятельности и применения технологий ИИ. Необходимо ограничивать результаты использования таких систем границами первичной обработки данных и выдачи рекомендаций человеку, принимающему решения. Кроме того, от подобных систем потребуются изложение примененных методов и логики выработки предлагаемых решений.

5. «Короткая» память систем ИИ. При непрерывном обучении новая информация перезаписывает более раннюю; этот эффект называют катастрофическим забыванием.

6. Самоактуализация является вопросом ценностей и смыслов. Для основы структурирования ограничений развития технологий технических систем ИИ использована иерархическая модель потребностей человека, представляющая собой упрощенное изложение идей американского ученого А. Маслоу. Однако может ли машина, как и человек, иметь стремление к развитию своих личностных возможностей? И не является ли это вымыслом фантастов? Для обсуждения этих вопросов обратимся к ключевым понятиям синергетики – междисциплинарного направления, в котором результаты исследований по теории лазеров использованы как идейная основа выявления общих закономерностей функционирования живых и неживых систем. Синергетика использует три ключевые идеи: самоорганизация, открытость, нелинейность. В результате самоорганизации через этап хаоса возникает новый порядок или новые структуры. Системы, в которых возможна самоорганизация, должны обладать высоким уровнем сложности взаимодействующих элементов, обмениваться с окружающим пространством веществом, энергией, энтропией (информацией). Даже поверхностный взгляд на ИНС с позиций синергетики позволяет априорно утверждать, что эти системы должны обладать возможностью к самоорганизации, которая произойдет через этап динамического

хаоса. Важно отметить, что системы ИИ являются сложными программно-техническими реализациями нелинейных алгоритмов, им абсолютно чужды совесть и эмпатия, они функционируют в собственной сконструированной реальности, интернально демонстрируют способность к самоорганизации. Системное понимание возможностей экстернальной самоорганизации вероятно на принципах синергетики.

Перспективы технологий искусственного интеллекта

Проведем структуризацию перспектив развития ИИ по представленному выше иерархическому принципу с добавлением еще одного уровня «пирамиды» – энергии.

1. *Энергетика.* Формирование индустриального общества стало возможным после первой промышленной революции, которая была основана на инновациях, использовании новых источников энергии – воды (текстильная промышленность), пара (паровые двигатели), каменноугольного кокса (на предприятиях), газа (освещение). Для второй промышленной революции характерна смена основных источников энергии, в промышленности и экономике активно используются электричество и энергия углеводородов (двигатели внутреннего сгорания). Для третьей промышленной революции характерно удешевление энергии в результате использования энергии атома. Для разворачивающейся в настоящее время четвертой промышленной революции предполагается возможность использования альтернативных возобновляемых источников энергии. Однозначно можно говорить о возможностях качественного изменения технологий планирования производства энергии. Применение технологий ИИ позволяет получить точность прогнозирования электропотребления на уровне 96–97%, что дает возможность повысить эффективность энергосбытовых компаний. Например, программа развития электросетевой компании «Россети» [19] предусматривает переход к интеллектуальным сетям типа Smart Grid, которые обладают качественными новыми характеристиками надежности, эффективности, доступности и управляемости. Использование энергии является основой для промышленного производства.

2. *Промышленность.* Инновационные технологии использования энергетических ресурсов позволяют достичь роста производительности труда в десятки раз по сравнению с предыдущим технологическим укладом. Промышленность является аналогом физической среды индустриального общества. Предприятия и организации



во всех отраслях увеличивают свои инвестиции в ИИ с целью сохранения и роста конкурентоспособности.

В 2012 г. произошла третья волна крупных инвестиций в ИИ. Объем инвестиций в технологии ИИ составил около 346,4 млрд долл. и за восемь лет вырос в 26 раз. ИИ присутствует почти в каждом из продуктов экосистемы Google (Search, Cloud, Gaming). Согласно отчетам консалтинговых компаний, количество транзакций, созданных различными информационными системами, с 2019 до 2025 г. вырастет с 10 до 200 млрд долл. [20], к 2027 г. – до 327 млрд долл., максимальные темпы роста применения ИИ будут приходиться (в единицах CAGR – совокупный среднегодовой темп роста): производство – 31,1%; логистика – 24,7%; здравоохранение – 22,0%; финансовый сектор – 21,0%.

Развитие постиндустриального Общества 5.0 в условиях формирования Индустрии 4.0 привело к высокой степени автоматизации и цифровизации промышленности, что открывает возможности для развития сферы ИИ. Консалтинговые компании предполагают, что на основе накопленной в базах данных промышленных предприятий информации именно промышленность станет одним из ключевых драйверов развития ИИ.

Ведущая роль в развитии технологий ИИ в таких странах, как США, Великобритания, странах ЕС принадлежат специализирующимся в области ИИ крупным корпорациям. В России большое значение будет иметь нефтегазовый бизнес. В 2021 г. ПАО «Газпром нефть» выступило в качестве учредителя Ассоциации «Искусственный интеллект в промышленности».

3. *Безопасность.* Активно внедряются технологии киберфизических систем и метавселенной, в которых обеспечена совместная работа человека и роботов, отчасти основанная на использовании коллаборативных роботов (коботы). Технологии разработки коботов подразумевают, что они не могут причинить вред человеку или окружающему физическому пространству. Использование цифровых двойников в киберфизических системах предполагается на всех этапах жизненного цикла продукции и управления производством.

4. *Системы коммуникаций.* Одним из ограничений систем ИИ является специфика функционирования в собственном сконструированном пространстве. В настоящее время ожидается, что искусственно созданные ограниченные проекции реальностей будут объединены в метавселенную, ключевая сила которой заключается

в способности обеспечивать целостный опыт погружения в виртуальную среду благодаря повсеместному и постоянному подключению. В промышленной метавселенной могут содержаться цифровые двойники реальных объектов, имитационные модели распределенных систем, информационные интерактивные системы управления производством.

Можно выделить следующие аспекты использования промышленной метавселенной:

- исследования, разработки, проектирование, моделирование и тестирование продуктов и услуг методами технологий ИИ;
- продвижение бренда методами взаимодействия с клиентами при проведении интерактивных виртуальных мероприятий презентации продуктов и услуг, демонстрации всех этапов производства;
- развитие взаимодействия с цифровыми двойниками клиентов для создания обратной связи и улучшения качества продукции в реальном мире;
- платежи с использованием технологий блокчейна, криптовалют и NFT.

ИИ станет базовым элементом в архитектуре современных кастомизированных производственных систем. Преимущества таких систем – улучшенная производственная эффективность, облегченное профилактическое обслуживание, развитые умные цепочки поставок [21].

Создание кастомизированных производств и метавселенной потребует не только развития и внедрения технологий ИИ, но и преодоления «цифрового ментального разрыва», формирования возможности принятия участниками производственных процессов и их клиентов взаимодействия в виртуальной среде.

5. *Достижение успеха и высокой оценки, признание.* Быстрое развитие и внедрение технологий ИИ требует изменения регулирования прав на результаты создания цифровых услуг, интеллектуальной деятельности, регулирования правомерности доступа технологий ИИ к личным данным и цифровым объектам, признания прав на виртуальные объекты цифровой экосистемы, понимания и правовой легитимности принятия решений на основе работы технологий ИИ, например, в медицинских системах диагностики и др.

Для технологического и промышленного развития России необходимы инновации и реинжиниринг [22]. Современный реинжиниринг, в свою очередь, нуждается в технологиях ИИ. Сокращение импорта микроэлектроники для ИИ



создает угрозу разрыва производственных цепочек. Решением проблемы должен стать запуск программ обратного инжиниринга аппаратного и программного обеспечения. При этом проводится исследование некоторого устройства или программы и сопроводительной документации на него для выявления его принципов работы и воспроизводства аналога [23]. В то же время технологии реинжиниринга потребуют пересмотреть законодательство в области охраны интеллектуального права и, возможно, изменить установки и общественный взгляд на копирование как явление [24]. ИИ сможет стать основой для реинжиниринга и импортозамещения, только если будет обеспечен российской продукцией микроэлектроники.

6. **Самоактуализация.** Формирование Общества 5.0 в результате четвертой промышленной революции потребует трансформации понимания роли технологий ИИ, приведет к созданию специфического интеллектуального пространства, взаимодействующего с метавселенной. Реализация сопутствующих процессов потребует инновационных методов управления и мониторинга для обеспечения позитивных результатов взаимодействия.

Угрозы развития и использования систем искусственного интеллекта

Технологии ИИ способны творить настоящую магию и имеют большие перспективы по изменению жизни людей. Однако реализация этих технологий для человечества, возможно, будет иметь угрожающе высокую цену.

1. **Энергетика.** Первый круг проблем – экологические. В последние 15–20 лет ряд общественных организаций представляют «глобальное потепление» как главную угрозу человечеству, возможно, более опасную, чем даже ядерная катастрофа. Решение задачи снижения углеродного следа корпорации стремятся переложить на плечи отдельных потребителей, призывая рассчитать и уменьшить личный углеродный след. Активно внедряются попытки форсировать переход от традиционных источников энергии к возобновляемым. Проблемы климата и экологии используются как инструменты политической игры, направленной на получение экономических выгод для крупных корпораций и отдельных государств. В качестве примера можно привести решение Европейского союза ввести сбор на ввоз продукции с высоким углеродным следом, которой являются минеральное сырье, черные и цветные металлы и др. При этом из широкого обсуждения научной общественностью выводятся такие

аспекты, как первичная доля углекислого газа в атмосфере или температура на Земле, или как велик углеродный след систем ИИ.

Авторы работы [25] провели оценку жизненного цикла для обучения распространенным крупным моделям ИИ, таким как Transformer, ELMo, BERT и GPT. Они обнаружили, что этот процесс может выделять более 626 000 фунтов (283 949 кг) эквивалента углекислого газа. Этот углеродный след в пять раз превышает след жизненного цикла среднего американского автомобиля (производство автомобиля, утилизацию, добычу нефти, транспортировку, производство топлива, выбросы от сжигания топлива). Отметим, что тренировка одной модели – это минимальный объем требуемой работы. Авторы полагают, что реальные показатели вредных выбросов в атмосферу от работающей модели ИИ окажутся в несколько раз больше.

Microsoft, Google и Amazon – крупнейшие облачные компании США – все имеют обязательства с отрицательным или нейтральным выбросом углерода. В заявлении Google говорится, что она стремится к 2030 г. достичь нулевого уровня выбросов во всех своих операциях, поставив перед собой цель полностью использовать свои офисы и центры обработки данных на безуглеродной энергии.

Лаборатория искусственного интеллекта Сбербанка в партнерстве с Институтом искусственного интеллекта AIRI создали open-source-библиотеку Eco2AI для оценки эквивалентного углеродного следа в зависимости от электроэнергии, затраченной на обучение моделей машинного обучения [26].

2. **Промышленность.** Отток инвестиций в отрасли. Исследователи отмечают ряд глобальных предпосылок, способных ввести рынок ИИ в состояние «зимы», однако полагают это маловероятным, возможным только в ряде сегментов. Для России существует два специфических направления, способных заморозить развитие в области технологий ИИ и, как следствие, создать условия для непреодолимого отставания от стран, владеющих технологиями ИИ. Первое – комплектующие микропроцессорной техники. Вследствие санкций сокращается/нарушаются цепочки поставок компонентной базы полупроводниковых устройств. Особенно остро воздействуют санкции на поставки продукции двойного назначения – чипы и полупроводники. По оценке Минэкономразвития РФ, объем импорта полупроводниковой продукции, попадающего под санкции, составит около 470 млн долларов, это около четверти всего



импорта продукции данного типа. Второе направление – это информационная изоляция из-за ограничения двустороннего обмена научно-технической информацией. Особенность современной науки в том, что прорывные исследования и разработки возникают преимущественно в обширных мультидисциплинарных и междотраслевых суперкластерах. Встает вопрос об отечественных разработках в области технологий ИИ, включая производство физической реализации технических систем. Однако такие разработки требуют больших временных и финансовых затрат, для которых, в свою очередь, требуется ИИ.

3. *Безопасность.* В области безопасности существуют риски противоправных операций с личными данными людей в киберфизических системах, возможность манипулирования системами жизнеобеспечения и безопасности со стороны ИИ. По мнению специалистов Artificial Intelligence Index Report, машинное мышление может начать локальные вооруженные конфликты с человеком с использованием автономных объектов (БПЛА, военные образцы робототехники), с учетом темпов роста машинного IQ вероятность такого исхода составляет, по оценке военных аналитиков, 30–35% [27].

4. *Социальные связи:* Если технологии ИИ будут воспроизводить модели поведения человека, то есть вероятность (20–24%), что отдельные ядра ИИ могут организовать свой цифровой рынок с правилами конкурентной борьбы с хакерскими атаками друг на друга.

5. *Признание.* В марте 2023 г. некоммерческая организация Future of Life опубликовала открытое письмо «Приостановить гигантские эксперименты с искусственным интеллектом», в котором ее глава Илон Маск и почти 1300 исследователей ИИ призвали все лаборатории искусственного интеллекта немедленно приостановить обучение систем искусственного интеллекта, более мощных, чем GPT-4, так как человечество не может позволить избранным техническим лидерам рисковать потерей контроля над нашей цивилизацией. Мощные системы искусственного интеллекта должны разрабатываться только тогда, когда человечество/люди будут уверены, что их эффект будет положительным, а их риски – управляемыми.

Также существуют риски введения такого моратория. Мораторий на разработку ИИ может привести к ряду непредвиденных последствий, которые в конечном итоге могут оказаться контрпродуктивными. Одним из возможных результатов является перенос исследований ИИ в теневую среду, где исследователи могут чувствовать себя

обязанными продолжать свою работу без контроля со стороны регулирующих органов или более широкого научного сообщества. Такие тайные усилия могут привести к отсутствию прозрачности и надзора, что усугубит риски, связанные с ИИ, а не уменьшит их.

6. *Самоактуализация.* Системы ИИ уже достаточно хорошо решают задачи программирования на разных языках. Через системы распознавания речи возможна эволюция естественных и искусственных языков в некую единую языковую систему. А когда меняется язык, меняется и мышление.

Заключение

Таким образом, предложенная в настоящей работе структура направлена на постановку проблемы разработки методологических и аналитических инструментов для оценки показателей развития систем ИИ. Предложен подход к разработке структурирования оценок показателей внедрения технологий ИИ.

В результате проведенного исследования получена структурированная система взаимосвязанных ограничений – перспектив – угроз систем ИИ. Такая структура позволяет строить взаимосвязанные срезы показателей развития систем ИИ. Проведение ее анализа возможно методами математического моделирования. Для этого необходимо собрать данные и представить их в формате причинно-следственных событий, что позволит связать разнородные показатели. В результате будет формализована постановка задачи анализа и прогноза оценки возможностей развития технологий ИИ. Для этого требуется формирование моделей и алгоритмов проверки выполнимости плана воздействий с использованием математических моделей, позволяющих определить вероятность и нарушения исполнения этого плана при неблагоприятном стечении событий. С учетом концепции Стратегии научно-технологического развития Российской Федерации это позволит провести разработку и обоснование методологии оценки и контроля технологий ИИ математическими методами в дальнейшем.

Список литературы

1. Искусственный интеллект // Министерство цифрового развития, связи и массовых коммуникаций. URL: <https://digital.gov.ru/ru/activity/directions/1046/> (дата обращения: 24.03.2023).
2. АНО «Цифровая экономика» изучила лучшие практики внедрения ИИ в российскую обрабатывающую промышленность. URL: <https://data-economy.ru/>



- news/tpost/zfbdoavnco1-ano-tsifrovaya-ekonomika-izuchila-luchsh (дата обращения: 24.03.2023).
3. The Socio-Economic Impact of AI in Healthcare. URL: https://www.medtecheurope.org/wp-content/uploads/2020/10/mte-ai_impact-in-healthcare_oct2020_report.pdf (дата обращения: 24.03.2023).
 4. Haseeb M., Mihardjo L. W., Gill A. R., Jermittiparsert K. Economic impact of artificial intelligence: New look for the macroeconomic assessment in Asia-Pacific region // International Journal of Computational Intelligence Systems. 2019. Vol. 12, iss. 2. P. 1295–1320. <https://doi.org/10.2991/ijcis.d.191025.001>
 5. Van Roy V. AI Watch – National strategies on Artificial Intelligence: A European perspective in 2019 // EUR 30102 EN, Publications Office of the European Union, Luxembourg, <https://doi.org/10.2760/602843>, JRC119974
 6. Review of Artificial Intelligence and Machine Learning Technologies: Classification, Restrictions, Opportunities and Challenges // Mathematics, MDPI. 2022. Vol. 10, iss. 15. P. 1–25. <https://dx.doi.org/10.3390/math10152552>
 7. Аверкин А. Н., Гаазе-Панопорт М. Г., Поснелов Д. А. Толковый словарь по искусственному интеллекту. М. : Радио и связь, 1992. 256 с.
 8. Ивахненко А. Г. Самообучающиеся системы с положительными обратными связями. Киев : Изд-во АН УССР, 1963. 330 с.
 9. Вешнева И. В., Сингаулин Р. А. Разработка информационно-образовательных комплексов системы дистанционного обучения с обратной связью на основе фотограмметрических методов и статусных функций. Часть 2 // Вестник Саратовского государственного технического университета. 2015. Т. 3, № 1 (80). С. 125–132. EDN: VKUGQB
 10. Вешнева И. В. Математические модели в системе управления качеством высшего образования с использованием методов нечеткой логики. Саратов : Саратовский источник, 2010. 186 с.
 11. Bolshakov A. A., Veshneva I. V. Assessment of the Effectiveness of Decision Support in the Application of the Information System for Monitoring the Process of Forming Competences Based on Status Functions // 2018 International Conference on Actual Problems of Electron Devices Engineering (APEDE). Saratov, Russia, 2018. P. 75–82. <https://doi.org/10.1109/APEDE.2018.8542462>
 12. Заставь машину думать: как развивают искусственный интеллект у роботов // РБК. URL: <https://trends.rbc.ru/trends/innovation/5d6520ea9a79474acef15b63> (дата обращения: 24.03.2023)
 13. Islas-Cota E., Gutierrez-Garcia J. O., Acosta C. O., Rodríguez L. F. A systematic review of intelligent assistants // Future Generation Computer Systems. 2022. Vol. 128. P. 45–62. <https://doi.org/10.1016/j.future.2021.09.035>
 14. Krinkin K., Shichkina Y., Ignatyev A. Co-evolutionary hybrid intelligence // 2021 5th Scientific School Dynamics of Complex Networks and their Applications (DCNA). Kaliningrad, Russian Federation. 2021. P. 112–115, <https://doi.org/10.1109/DCNA53427.2021.9587002>
 15. Созданный в США робот пообещал уничтожить человечество. URL: <https://rg.ru/2016/03/22/robot-roobeshchal-unichtozhit-chelovechestvo.html> (дата обращения: 24.03.2023).
 16. Veshneva I. V., Bolshakov A. A. Construction of Digital Twins of Socio-Economic Systems Using Mathematical Models Based on Status Functions // Society 5.0: Human-Centered Society Challenges and Solutions. Cham : Springer, 2022. P. 129–141. https://doi.org/10.1007/978-3-030-95112-2_11
 17. Bolshakov A., Veshneva I., Lushin D. Mathematical Model of Integration of Cyber-Physical Systems for Solving Problems of Increasing the Competitiveness of the Regions of the Russian Federation // Society 5.0: Cyberspace for Advanced Human-Centered Society. Cham : Springer Verlag, 2021. P. 129–139. https://doi.org/10.1007/978-3-030-63563-3_11
 18. Veshneva I., Chernyshova G., Bolshakov A. Regional Competitiveness Research Based on Digital Models Using Kolmogorov-Chapman Equations // Society 5.0: Cyberspace for Advanced Human-Centered Society. Cham : Springer Verlag, 2021. P. 141–154. https://doi.org/10.1007/978-3-030-63563-3_12
 19. Перспективы развития // РОССЕТИ. Стратегический отчет. URL: <http://rosseti14.kerrygun.ru/strategic-report/prospects-of-development> (дата обращения: 24.03.2023).
 20. Investment in Artificial Intelligence Solutions Will Accelerate as Businesses Seek Insights, Efficiency, and Innovation, According to a New IDC Spending Guide // International Data Corporation. 2021. URL: <https://idc.com/getdoc.jsp?containerId=prUS48191221> (дата обращения: 24.03.2023).
 21. Княгинин В. Н., Липецкая М. С., Санатов Д. В., Васеев И. Е., Годунова Е. А., Семенова М. А., Харитонов М. А., Холоднова Е. М. Искусственный интеллект в промышленности. Экспертно-аналитический доклад. URL: https://csr-nw.ru/upload/iblock/3db/Доклад%20по%20ИИ%20в%20промышленности_финал.pdf (дата обращения: 24.03.2023).
 22. Чернышова Г. Ю., Вешнева И. В., Роках Г. Е. Моделирование динамики рисков региональной конкурентоспособности // Известия Саратовского университета. Новая серия. Серия: Экономика. Управление. Право. 2022. Т. 22, вып. 1. С. 42–52. <https://doi.org/10.18500/1994-2540-2022-22-1-42-52>
 23. Samuelson P., Scotchmer S. The Law and Economics of Reverse Engineering // The Yale Law Journal. 2002. Vol. 111, no. 7. P. 1575–1663. <https://doi.org/10.2307/797533>
 24. Shenkar O., Oded Shenkar, Yadong Luo. International business. Hoboken, N. J. : Wiley, 2004. 513, [217] p.
 25. Strubell E., Ganesh A., McCallum A. Energy and Policy Considerations for Deep Learning in NLP // Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics. Florence, Italy, 2019. P. 3645–3650. URL: <https://arxiv.org/pdf/1906.02243.pdf> (дата обращения: 24.03.2023).



26. Сбер представил открытый инструмент для оценки углеродного следа моделей ИИ // Lenta.Ru. URL: <https://lenta.ru/news/2022/06/17/openinstr/> (дата обращения: 24.03.2023).
27. Широковских С. А. Возможности, ограничения и вероятные угрозы устойчивому развитию высокотехнологических компаний азиатского региона вследствие выхода из-под контроля искусственного интеллекта и нейросетей // Экономика Центральной Азии. 2020. Т. 4, № 4. С. 385–394. <https://doi.org/10.18334/asia.4.4.111626>
- ### Reference
1. Artificial intelligence. *Ministerstvo tsifrovogo razvitiya, svyazi i massovykh kommunikatsiy* (Ministry of Digital Development, Communications and Mass Communications). Available at: <https://digital.gov.ru/ru/activity/directions/1046/> (accessed March 24, 2023) (in Russian).
 2. ANO «Tsifrovaya ekonomika» izuchila luchshiyе praktiky vnedreniya II v rossiyskuyu obrabatyvayushchuyu promyshlennost' (ANO Digital Economy studied the best practices for introducing AI into the Russian manufacturing industry). Available at: <https://data-economy.ru/news/tpost/zfbdoavnco1-ano-tsifrovaya-ekonomika-izuchila-luchsh> (accessed March 24, 2023) (in Russian).
 3. *The Socio-Economic Impact of AI in Healthcare*. Available at: https://www.medtecheurope.org/wp-content/uploads/2020/10/mte-ai_impact-in-healthcare_oct2020_report.pdf (accessed March 24, 2023).
 4. Haseeb M., Mihardjo L. W., Gill A. R., Jermstipparsert K. Economic impact of artificial intelligence: New look for the macroeconomic assessment in Asia-Pacific region. *International Journal of Computational Intelligence Systems*, 2019, vol. 12, iss. 2, pp. 1295–1320. <https://doi.org/10.2991/ijcis.d.191025.001>
 5. Van Roy V. AI Watch – National strategies on Artificial Intelligence: A European perspective in 2019. *EUR 30102 EN, Publications Office of the European Union, Luxembourg*. <https://doi.org/10.2760/602843>, JRC119974
 6. Review of Artificial Intelligence and Machine Learning Technologies: Classification, Restrictions, Opportunities and Challenges. *Mathematics, MDPI*, 2022, vol. 10, iss. 15, pp. 1–25. <https://dx.doi.org/10.3390/math10152552>
 7. Averkin A. N., Gaaze-Rapoport M. G., Pospelov D. A. *Tolkovyy slovar' po iskusstvennomu intellektu* [Explanatory dictionary on artificial intelligence]. Moscow, Radio i svyaz', 1992. 256 p. (in Russian).
 8. Ivahnenko A. G. *Samoobuchayushchiesya sistemy s polozhitel'nymi obratnymi svyazyami* [Self-learning systems with positive feedback]. Kiev, Academy of Sciences of the Ukrainian SSR Publ., 1963. 330 p. (in Russian).
 9. Veshneva I. V., Singatulin R. A. Development of information measuring systems for distance learning complexes with feedback based on photogrammetric methods and status functions. Part 2. *Vestnik of Saratov State Socio-Economic University*, 2015, vol. 3, no. 1 (80), pp. 125–132 (in Russian). EDN: VKUGQB
 10. Veshneva I. V. *Matematicheskie modeli v sisteme upravleniya kachestvom vysshego obrazovaniya s ispol'zovaniem metodov nechetkoj logiki* [Mathematical models in the quality management system of higher education using fuzzy logic methods]. Saratov, Saratovskiy istochnik, 2010. 186 p. (in Russian).
 11. Bolshakov A. A., Veshneva I. V. Assessment of the Effectiveness of Decision Support in the Application of the Information System for Monitoring the Process of Forming Competences Based on Status Functions. *2018 International Conference on Actual Problems of Electron Devices Engineering (APEDE)*. Saratov, 2018, pp. 75–82. <https://doi.org/10.1109/APEDE.2018.8542462>
 12. Make the machine think: How artificial intelligence is developed in robots. *RBK*. Available at: <https://trends.rbc.ru/trends/innovation/5d6520ea9a79474acef15b63> (accessed March 24, 2023) (in Russian).
 13. Islas-Cota E., Gutierrez-Garcia J. O., Acosta C. O., Rodríguez L. F. A systematic review of intelligent assistants. *Future Generation Computer Systems*, 2022, vol. 128, pp. 45–62. <https://doi.org/10.1016/j.future.2021.09.035>
 14. Krinkin K., Shichkina Y., Ignatyev A. Co-evolutionary hybrid intelligence. *2021 5th Scientific School Dynamics of Complex Networks and their Applications (DCNA)*. Kaliningrad, Russian Federation, 2021, pp. 112–115. <https://doi.org/10.1109/DCNA53427.2021.9587002>
 15. *Sozdannyy v SShA robot poobeshchal unichtozhit' chelovechestvo* (A robot created in the USA promised to destroy humanity). Available at: <https://rg.ru/2016/03/22/robot-poobeshchal-unichtozhit-chelovechestvo.html> (accessed March 24, 2023) (in Russian).
 16. Veshneva I. V., Bolshakov A. A. Construction of Digital Twins of Socio-Economic Systems Using Mathematical Models Based on Status Functions. In: *Society 5.0: Human-Centered Society Challenges and Solutions*. Cham, Springer, 2022, pp. 129–141. https://doi.org/10.1007/978-3-030-95112-2_11
 17. Bolshakov A. Veshneva I., Lushin D. Mathematical Model of Integration of Cyber-Physical Systems for Solving Problems of Increasing the Competitiveness of the Regions of the Russian Federation. In: *Society 5.0: Cyberspace for Advanced Human-Centered Society*. Cham, Springer Verlag, 2021, pp. 129–139. https://doi.org/10.1007/978-3-030-63563-3_11
 18. Veshneva I. Chernyshova G., Bolshakov A. Regional Competitiveness Research Based on Digital Models Using Kolmogorov-Chapman Equations. In: *Society 5.0: Cyberspace for Advanced Human-Centered Society*. Cham, Springer Verlag, 2021, pp. 141–154. https://doi.org/10.1007/978-3-030-63563-3_12
 19. Development prospects. *ROSSETI. Strategicheskij otchet* (ROSSETI. Strategic report). Available at: <http://rosseti14.kerrygun.ru/strategic-report/prospects-of-development> (accessed March 24, 2023) (in Russian).
 20. Investment in Artificial Intelligence Solutions Will Accelerate as Businesses Seek Insights, Efficiency, and Innovation, According to a New IDC Spending Guide.



- International Data Corporation*. 2021. Available at: <https://idc.com/getdoc.jsp?containerId=prUS48191221> (accessed March 24, 2023).
21. Knyagin V. N., Lipetskaya M. S., Sanatov D. V., Vaseev I. E., Godunova E. A., Semenova M. A., Kharitonov M. A., Kholodnova E. M. *Iskusstvennyy intellekt v promyshlennosti. Ekspertno-analiticheskiy doklad* (Artificial intelligence in industry. Expert analytical report). Available at: https://csr-nw.ru/upload/iblock/3db/Доклад%20по%20ИИ%20в%20промышленности_финал.pdf (accessed March 24, 2023) (in Russian).
22. Chernyshova G. Yu., Veshneva I. V., Rokakh G. E. Modeling the dynamics of regional competitiveness risks. *Izvestiya of Saratov University. Economics. Management. Law*, 2022, vol. 22, iss. 1, pp. 42–52 (in Russian). <https://doi.org/10.18500/1994-2540-2022-22-1-42-52>
23. Samuelson P., Scotchmer S. The Law and Economics of Reverse Engineering *The Yale Law Journal*, 2002, vol. 111, no. 7, pp. 1575–1663. <https://doi.org/10.2307/797533>
24. Shenkar O., Oded Shenkar, Yadong Luo. *International business*. Hoboken, N.J., Wiley, 2004. 513, [217] p.
25. Strubell E., Ganesh A., McCallum A. Energy and Policy Considerations for Deep Learning in NLP. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy, 2019, pp. 3645–3650. Available at: <https://arxiv.org/pdf/1906.02243.pdf> (accessed March 24, 2023).
26. Sber presented an open tool for assessing the carbon footprint of AI models. *Lenta.Ru*. Available at: <https://lenta.ru/news/2022/06/17/openinstr/> (accessed March 24, 2023) (in Russian).
27. Shirokovskih S.A. Opportunities, limitations and probable threats to the sustainable development of high-tech companies in the Asian region due to the out-of-control of artificial intelligence and neural networks. *Ekonomika Tsentralnoy Azii*, 2020, vol. 4, iss. 4, pp. 385–394 (in Russian). <https://doi.org/10.18334/asia.4.4.111626>

Поступила в редакцию 01.05.2023; одобрена после рецензирования 08.10.2023; принята к публикации 10.10.2023
The article was submitted 01.05.2023; approved after reviewing 08.10.2023; accepted for publication 10.10.2023