

Научная статья

УДК 330.47, 338.36, 339.9

JEL: C18, C45, F21, F47

<https://doi.org/10.18184/2079-4665.2024.15.2.228-247>

Искусственный интеллект: дихотомия развития технологии

Щербаков Геннадий Анатольевич¹

¹ Финансовый университет при Правительстве Российской Федерации; Москва, Россия

¹ g.shcherbakov@mail.ru, <https://orcid.org/0000-0002-3298-1638>

Аннотация

Цель: определение возможностей и потенциальных рисков внедрения технологий искусственного интеллекта на основе анализа доминирующих тенденций современного развития высокотехнологичного сектора глобальной экономики.

Методы. В работе применены системный подход, метод классификации и систематизации, метод оценок и научных абстракций, методы информационного анализа и синтеза, метод визуализации данных.

Результаты работы. В статье рассмотрены экономические и технико-технологические аспекты построения интеллектуальных систем, установлены тенденции их развития в направлении многозадачности и автономности выполняемых функций, определены возможности и угрозы, возникающие в результате развития данного направления научно-технического прогресса. Исследование выявило двойственный характер дальнейшего совершенствования систем искусственного интеллекта, отрывающего широкие перспективы применения в различных сферах жизнедеятельности человека, однако способного одновременно генерировать риски вследствие бесконтрольного внедрения данных технологий в общественное пространство.

Выводы. Развитие интеллектуальных технологий в текущем периоде имеет устойчивую тенденцию к усложнению решаемых задач и повышению автономности искусственного интеллекта, обеспечиваемую наращиванием его вычислительных способностей, которые по ряду параметров (классификация изображений, содержание текстов) превосходят возможности человеческого разума. При этом прослеживается стремление разработчиков к постепенной передаче машинам ключевых компетенций в области принятия решений, до настоящего времени оставшихся в исключительной ответственности человека. Указанная тенденция открывает широкие перспективы применения «умных» машин в общественной жизни, однако одновременно генерирует комплекс реальных и гипотетических рисков, доминирующим из которых выступает утрата человеком субъектности по отношению к окружающим процессам. Сформировавшаяся дихотомия технологического развития требует серьезного переосмысления существующего подхода к созданию систем искусственного интеллекта в целях минимизации возникновения этических рисков и угроз жизнедеятельности человека.

Ключевые слова: мировое технологическое развитие, цифровизация, искусственный интеллект, экономико-математическое моделирование, генеративные системы, мультимодальные модели, технологическая этика

Конфликт интересов. Автор заявляет об отсутствии конфликта интересов.

Для цитирования: Щербаков Г. А. Искусственный интеллект: дихотомия развития технологии // МИР (Модернизация. Инновации. Развитие). 2024. Т. 15. № 2. С. 228–247

EDN: <https://elibrary.ru/incmkf>. <https://doi.org/10.18184/2079-4665.2024.15.2.228-247>

© Щербаков Г. А., 2024



Контент доступен под лицензией Creative Commons Attribution 4.0 License.
The content is available under Creative Commons Attribution 4.0 License.

Artificial Intelligence: the dichotomy of technology development

Gennady A. Shcherbakov¹

¹ Financial University under the Government of the Russian Federation; Moscow, Russia

¹g.shcherbakov@mail.ru, <https://orcid.org/0000-0002-3298-1638>

Abstract

Purpose: identification of opportunities and potential risks in the implementation of artificial intelligence technologies on the basis of the dominant trends analyzing in the current high-tech sector development of the global economy.

Methods: the paper uses the system approach, the method of classification and systematization, the method of estimations and scientific abstractions, the methods of information analysis and synthesis, the method of data visualization.

Results: the article considers economic and technical-technological aspects of building intelligent systems, establishes trends of their development in the direction of multitasking and autonomy of performed functions, identifies opportunities and threats arising from the development of this direction of scientific and technological progress. The study reveals the dual nature of further improvement of artificial intelligence systems, which offers broad prospects of application in various spheres of human activity, but can simultaneously generate risks due to the uncontrolled introduction of these technologies into the public space.

Conclusions and Relevance: the development of intellectual technologies in the current period has a steady tendency towards increasing complexity of the tasks to be solved and increasing autonomy of artificial intelligence, provided by increasing its computational capabilities, which in a number of parameters (image classification, text content) exceed the capabilities of the human mind. At the same time, there is an aspiration of developers to gradually transfer to machines the key competences in the field of decision-making, which until now have remained in the exclusive competence of humans. This trend opens up broad prospects for the application of "smart" machines in public life, but at the same time generates a set of real and hypothetical risks, the dominant of which is the loss of human subjectivity in relation to the surrounding processes. The emerging dichotomy of technological development requires a serious rethinking of the existing approach to the creation and implementation of artificial intelligence systems in order to minimize the emergence of ethical risks and threats to human life.

Keywords: global technological development, digitalization, artificial intelligence, economic and mathematical modelling, generative systems, multimodal language models, technological ethics

Conflict of Interest. The author declares that there is no conflict of interest.

For citation: Shcherbakov G. A. Artificial Intelligence: the dichotomy of technology development. *MIR (Modernizatsiia. Innovatsii. Razvitie) = MIR (Modernization. Innovation. Research)*. 2024; 15(2):228–247. (In Russ.)

EDN: <https://elibrary.ru/incmkf>. <https://doi.org/10.18184/2079-4665.2024.15.2.228-247>

© Shcherbakov G. A., 2024

Введение

Текущий период отмечен нарастающим процессом внедрения цифровых технологий в различные сферы деятельности общества, где выполнение функций, находившихся ранее в сфере исключительной ответственности человека, теперь осуществляется с помощью технологических нововведений.

Замещение тех или иных действий человека операциями, выполняемыми технико-технологическими средствами, происходило практически на протяжении всего развития общества. Однако в прошлом технологии, направленные на замену человеческих функций машинами, носили, в основном, характер замещения физических усилий человека за счет манипулирования объектами с

помощью вспомогательных средств. Машины были задействованы для выполнения рутинных, повторяющихся задач, таких как, например, работа на ткацком станке, заменившая труд ручных ткачей. Такие технологии формировали вспомогательный инструментарий в жизни человека, служили средством активизации производственных процессов и не затрагивали субъектность индивида в отношении его деятельности. Данное направление научно-технического прогресса по времени простирается от изобретения самых простых машин до создания робототехники.

Однако к настоящему периоду технологии продвинулись до имитации сложных когнитивных процессов, включая математические расчеты,

распознавание речи и письменное выражение мыслей. Новые технологии теперь направлены на воспроизведение мыслительных процессов в области анализа и принятия решений, что, по сути, соответствует человеческой интеллектуальной деятельности. Столь быстрые темпы распространения подобных технологий порождают опасения, что уже в ближайшем будущем не останется областей человеческой когнитивной активности, которые не будут автоматизированы или воспроизведены в новых интеллектуальных технологиях.

Использование в различных областях общественной жизни технологических нововведений, обобщенно получивших название «искусственный интеллект», является объективным результатом углубления научно-технического прогресса, его распространения на новые сферы деятельности. Однако практически бесконтрольный характер развития новейших технологий, отсутствие установленных границ его приложения чревато рисками утраты индивидуумом субъектности по отношению к процессам, протекающим в окружающем мире.

Обзор литературы и исследований

Еще в 1640 г. Т. Гоббс, в труде «Человеческая природа» [1], а в 1647 г. Р. Декарт, в своем «Рассуждении о методе» [2], затронули тему воспроизведения машинами когнитивных способностей человека. А спустя три столетия, в середине XX в., английский математик А. Тьюринг представил концепцию «Baby Machine» – устройства, способного воспринимать и аккумулировать данные, а также обучаться по аналогии с обучением ребенка, но на основе информации, рассчитанной на взрослого человека. В тот же период (1956 г.) состоялся знаменитый Дартмутский научный семинар, где американский математик Дж. Маккарти впервые использовал понятие «искусственный интеллект» как научный термин.

Обсуждая проблемы создания искусственного интеллекта в русскоязычной аудитории, следует осознавать, что термин «искусственный интеллект» в русском языке не всегда точно отражает смысл английского выражения «artificial intelligence». Его лучше переводить как «искусственный разум» – хотя «intelligence» на английском действительно означает «интеллект», но для этого понятия в русском языке уже существует аналогичный термин. Интеллект – это способность мыслить, а разум – более высокий уровень мышления, включающий логическое и творческое мышление. В 1956 г., когда Дж. Маккарти впервые употребил понятие

«artificial intelligence», говоря о моделировании когнитивных способностей человека, он, скорее всего, имел в виду именно воспроизведение способностей человеческого разума.

За прошедший период разработано множество методов и алгоритмов создания искусственного интеллекта, таких как «байесовские методы, логистическая регрессия, метод опорных векторов, решающие деревья, ансамбли алгоритмов и др. Вместе с тем, наиболее продвинутым и успешным подходом является глубокое машинное обучение, которое использует глубокие нейронные сети» [3, с. 82].

Нейронные сети представляют собой «упрощенную модель организации нервной системы биологического организма, где головной мозг содержит в общей сложности около 65 млрд нейронов и 100 трлн синапсов. Эти сети представляют собой модели организации нервной системы, состоящие из нейронов и синапсов. Нейронные сети позволяют машинам обучаться методом проб и ошибок»¹.

Математическая модель искусственной нейронной сети, основанная на функции нейрона $Y = F(C)$ как порогового устройства, была предложена У. Мак-Каллоком и У. Питтсом еще в 1943 г.:

$$S = \sum_{i=1}^n X_i \cdot w_i \quad (1)$$

Несмотря на широкое распространение идеи создания искусственного интеллекта в 1950-х гг., последовавшие три десятилетия оказались для нее неудачными из-за ограниченных возможностей доступных вычислительных устройств. Однако ситуация начала меняться в середине 1990-х гг., и сегодня разработка технологий искусственного интеллекта стала важнейшим направлением развития научно-технического прогресса, преследующим целью улучшение жизнеобеспечения людей и облегчение их деятельности. Именно в текущий период появилась возможность работать с большими объемами данных и использовать более сложные алгоритмы машинного обучения, что привело к новому всплеску интереса к искусственному интеллекту и машинному обучению.

Несмотря на относительную новизну вопросов, связанных с развитием технологий искусственного интеллекта, указанная проблематика нашла широкое освещение в научной литературе. В большинстве случаев исследования посвящены технико-технологическим аспектам развития искусственного интеллекта. К ним относятся монографии и статьи авто-

¹ Щербаков Г.А. Искусственный интеллект: новая этика современного технологического развития // Мягкие измерения и вычисления. 2021. Т. 48. № 11. С. 44–53. EDN: <https://elibrary.ru/bivxuo>. <https://doi.org/10.36871/2618-9976.2021.11.002>

ров, охватывающих период от первой половины XX столетия до настоящего времени (например, [4–14]). Экономические аспекты развития искусственного интеллекта достаточно представлены в ряде научных работ (например, [15, 16]), а также в различных периодических обзорах².

В наименьшей степени исследованными остаются вопросы, касающиеся возможных рисков, способных возникнуть в результате создания и применения технологий искусственного интеллекта. Представленные публикации обладают несомненными достоинствами, однако зачастую содержат заметные упущения, заключающиеся в упрощении ситуации, в частности, в сведении исследования к «техническим», преодолимым рискам (например, вытеснение рабочей силы, потеря конфиденциальности, ограничение плюрализма и проч.), не несущим экзистенциальной угрозы для человека³. В них зачастую игнорируются этические вопросы, связанные с перспективами развития «генеративного» искусственного интеллекта, способного в обозримой перспективе достичь уровня человека не только в его когнитивных возможностях, но и в принятии управленческих решений. Тем не менее, научное сообщество постепенно смещает фокус внимания в сторону обсуждения этических проблем развития искусственного интеллекта (например, [17–23]).

Следует признать, что сфера «машинной этики», несмотря на существенную теоретическую работу, проделанную в области интеллектуальных технологий, до сих пор представляет собой проблемное поле и по степени изученности заметно отстает от других исследований в области высоких технологий. Это обстоятельство требует пересмотра существующего подхода, а также серьезного переосмысления самой парадигмы процесса создания систем искусственного интеллекта в целях минимизации возникновения этических рисков. В указанной логике проведено исследование, результаты которого изложены в представленной статье.

Материалы и методы

Исследование искусственного интеллекта является сложной междисциплинарной задачей, кото-

рая акцентируется в основном на практических результатах, а не на фундаментальных исследованиях. В фокусе настоящей статьи находятся показатели анализа развития технологии в период с 2013 г. до настоящего времени. Учитывая, что работы в области искусственного интеллекта можно рассматривать как ускоренный, гипердинамичный процесс, вовлекающий в свое развитие огромные финансовые, материальные и интеллектуальные ресурсы общества, 10-летний срок исследования не представляется слишком коротким для состоятельных обобщений и выводов.

Эмпирическая база работы представлена комплексом тематических материалов, публикуемых в периодических научных изданиях. В исследовании, поэтапно проведенном в формате сбора и анализа статистических и фактологических данных, активно использовались методы обработки информации, полученной как из первичных источников (базы данных национальных статистических учреждений, международных экономических организаций и др.) или периодических обзоров (например, ежегодные отчеты Стэнфордского университета «об индексе искусственного интеллекта»⁴), так и из научных работ зарубежных и отечественных авторов. Были применены системный подход, метод классификации и систематизации, метод оценок и научных абстракций, метод сравнений, методы информационного анализа и синтеза, метод визуализации данных.

Результаты исследования

Технологические аспекты искусственного интеллекта

Современный этап научно-технического прогресса, характеризуемый стремительным ростом вычислительных возможностей и существенным прогрессом в развитии методов «машинного» обучения, способствовал возникновению нового поколения крупноразмерных систем искусственного интеллекта, названных «генеративными моделями» с глубоким обучением.

Системы искусственного интеллекта в настоящее время развиваются на следующих направлени-

² The state of AI in 2022 – and a half decade in review // QuantumBlack by McKinsey. URL: <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review> (дата обращения: 23.02.2024); The AI index 2023 annual report // AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA. April 2023. URL: <https://hai.stanford.edu/research/ai-index-2023> (дата обращения: 10.02.2024); The AI Index 2024 Annual Report // AI Index Steering Committee, Institute for Human-Centered AI, Stanford University. Stanford, CA: April 2024. URL: https://aiindex.stanford.edu/wp-content/uploads/2024/04/HAI_AI-Index-Report-2024.pdf (дата обращения: 18.04.2024).

³ Thomas M. 12 risks and dangers of artificial intelligence (AI) // Built In. URL: <https://builtin.com/artificial-intelligence/risks-of-artificial-intelligence> (дата обращения: 18.02.2024); Marr B. The 15 biggest risks of artificial intelligence // Forbes. URL: <https://www.forbes.com/sites/bernardmarr/2023/06/02/the-15-biggest-risks-of-artificial-intelligence/?sh=38f6584d2706> (дата обращения: 15.02.2024).

⁴ The AI Index report // Stanford University. URL: <https://aiindex.stanford.edu/report/> (дата обращения: 18.04.2024).

ях: NL⁵-восприятие текста, NL-восприятие речи, NL-генерация, глубокое обучение, обучение с подкреплением, компьютерное зрение, распознавание лиц, цифровые двойники, автоматизация роботизированных процессов, рекомендательные системы, физическая робототехника, виртуальные агенты и др.

Новейшие технологии искусственного интеллекта становятся фактором активизации научно-технического прогресса в различных сферах деятельности общества. В частности, в 2022-2023 гг. они применялись для создания новых антител, увеличения результативности манипуляций с матрицами, синтеза водорода и проч. Так, компания DeepMind создала систему обучения с подкреплением для моделирования процесса управления плазмой ядерного синтеза в токамаке. Компания Google активно применяет свою языковую модель PaLM для совершенствования самой модели, а производитель графических процессоров nVidia применяет агента обучения с подкреплением для задач совершенствования конструкции микросхем, производимых для самих систем искусственного интеллекта.

Международная исследовательская группа развивает модель IndicNLG, предназначенную для NL-генерации 11-ти наречий индийского языка. Указанный проект открывает дополнительные перспективы для использования технологий искусственного интеллекта при решении задач генерации естественного языка в неанглийских языковых форматах.

В 2021 г. была создана генеративная система Copilot, способная конвертировать инженерное решение в программный код, а также осуществлять перевод между языками программирования. Она выступает в качестве средства поддержки разработчиков искусственного интеллекта, которые в случае возникновения проблемы кодирования формулируют ее на естественном языке, после чего модель Copilot генерирует решение в программном коде. Экспериментальным путем установлено, что инженеры, применявшие Copilot, затрачивали времени на выполнение поставленной задачи на 56% меньше, чем инженеры, работавшие без программной поддержки (71 / 161 мин.)⁶.

Большинство авторов крупномасштабных генеративных систем, отсчет существования которых

можно вести с момента создания GPT-2 в 2019 г., представляют исследовательские институты США (54,02%). Однако в 2020-2021 гг. о своих достижениях в построении крупномасштабных моделей искусственного интеллекта заявили исследователи из Англии (21,88%), Китая (8,04%), Канады (6,25%), Израиля (5,80%), ФРГ (3,12%) и Индии (0,89%).

В качестве успешных примеров интеллектуальных систем можно выделить выпущенные в 2022–2024 гг. в США модели PaLM 540B (создатель – компания Google), DALL-E 3 и GPT-4 Turbo, Sora (OpenAI), Grok (xAI), китайскую крупномасштабную систему GLM-130B (Университет Цинхуа), а также систему BLOOM, ставшую результатом работы международного коллектива из более чем 1000 исследователей, и др. (табл. 1).

Россия представляет более скромные результаты по сравнению с лидерами разработок искусственного интеллекта, однако также ведет успешные работы в указанной сфере. Прежде всего, они осуществляются усилиями отечественных компаний Сбер (GigaChat, Kandinsky 2.2) и Яндекс (YandexGPT, YandexART). Так, весной 2024 г. была представлена крупномасштабная языковая модель YandexGPT 3 Pro, относящаяся к третьему поколению «линейки» YandexGPT, которая способна (в сравнении с предыдущей версией, YandexGPT 2) эффективнее решать сложные задачи, лучше понимать контекст диалога, корректнее работать с фактами, давать более полные ответы, допускать меньше стилистических ошибок. Она, по утверждениям разработчиков, превосходит ChatGPT-3.5 Turbo в ответах на узкие классы запросов, важных для бизнеса⁷.

Современные успехи в совершенствовании аппаратного обеспечения и методов обучения нейронных сетей способствовали, как указывалось выше, появлению нового поколения крупномасштабных интеллектуальных систем, которые продемонстрировали заметное повышение точности во многих задачах, поставленных перед ними разработчиками. Однако эти улучшения точности зависят от наличия исключительно больших вычислительных ресурсов, которые требуют столь же значительного энергопотребления. В результате обучение и разработка этих моделей обходятся дорого, как с финансовой точки зрения, из-за стоимости оборудования и электроэнергии или времени облачных вычислений, так и с экологической точки зрения,

⁵Прим. Автора: NL – Natural Language, естественный язык.

⁶The AI index 2023 annual report // AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA. April 2023. URL: <https://hai.stanford.edu/research/ai-index-2023> (дата обращения: 10.02.2024).

⁷YandexGPT 3 // Яндекс. URL: <https://ya.ru/ai/gpt-3> (дата обращения: 31.05.2024).

Таблица 1

Наиболее известные крупноразмерные генеративные модели искусственного интеллекта

Table 1

The most famous large-scale generative models of artificial intelligence

Период создания	Наименование модели	Страна-разработчик
1	2	3
2019 год, 1 квартал	GPT-2	США
2019 год, 2 квартал	Grover-Mega	США
2019 год, 3 квартал	Megatron-LM (Original, 8.3B)	США
2019 год, 4 квартал	T5-11B	США
2019 год, 4 квартал	T5-3B	США
2020 год, 1 квартал	Meena	США
2020 год, 1 квартал	Turing NLG	США
2020 год, 2 квартал	GPT-3 175B (davinci)	США
2020 год, 3 квартал	ERNIE-GEN (large)	Китай
2021 год, 1 квартал	DALL-E	США
2021 год, 1 квартал	Wu Dao - Wen Yuan	Китай
2021 год, 1 квартал	GPT-Neo	-
2021 год, 2 квартал	PanGu-alpha	Китай
2021 год, 2 квартал	GPT-J-6B	США
2021 год, 2 квартал	HyperClova	Корея
2021 год, 2 квартал	CogView	Китай
2021 год, 2 квартал	Wu Dao 2.0	Китай
2021 год, 3 квартал	Ernie 4.0	Китай
2021 год, 3 квартал	Codex	США
2021 год, 3 квартал	Jurassic-1-Jumbo	Израиль
2021 год, 4 квартал	Megatron-Turing NLG 530B	США
2021 год, 4 квартал	Gopher	Англия
2022 год, 1 квартал	InstructGPT	США
2022 год, 1 квартал	AlphaCode	Англия
2022 год, 1 квартал	GPT-NeoX-20B	США, Англия, ФРГ, Индия
2022 год, 1 квартал	Chinchilla	Великобритания
2022 год, 1 квартал	PaLM (540B)	США
2022 год, 1 квартал	DALL-E 2	США
2022 год, 2 квартал	Stable Diffusion (LDM-KL-8-G)	Германия
2022 год, 2 квартал	OPT-175B	США
2022 год, 2 квартал	Jurassic-X	Израиль
2022 год, 2 квартал	Imagen	Канада
2022 год, 3 квартал	Minerva (540B)	США
2022 год, 3 квартал	GLM-130B	Китай
2022 год, 4 квартал	BLOOM	международная группа
2023 год, 1 квартал	Claude	США
2023 год, 1 квартал	GPT-4	США
2023 год, 2 квартал	GigaChat	Россия
2023 год, 3 квартал	Llama 2	США

Окончание таблицы 1

End of the table 1

1	2	3
2023 год, 3 квартал	Mistral 7B	Франция
2023 год, 4 квартал	Ernie 4.0	Китай
2023 год, 4 квартал	GPT-4 Turbo	США
2023 год, 4 квартал	Gemini	США
2024 год, 1 квартал	YandexGPT 3 Pro	Россия

Составлено автором по материалам: *The AI index 2023 annual report* // AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA. April 2023. URL: <https://hai.stanford.edu/research/ai-index-2023>; *The AI Index 2024 Annual Report* // AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA: April 2024. URL: https://aiindex.stanford.edu/wp-content/uploads/2024/04/HAI_AI-Index-Report-2024.pdf; YandexGPT 3 // Яндекс. URL: <https://ya.ru/ai/gpt-3>; GigaChat // Сбер. URL: <https://developers.sber.ru/gigachat/login>

Compiled by the author according to: *The AI index 2023 annual report*. AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA. April 2023. URL: <https://hai.stanford.edu/research/ai-index-2023>; *The AI Index 2024 Annual Report*. AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA: April 2024. URL: https://aiindex.stanford.edu/wp-content/uploads/2024/04/HAI_AI-Index-Report-2024.pdf; YandexGPT 3. URL: <https://ya.ru/ai/gpt-3>; GigaChat. Sber. URL: <https://developers.sber.ru/gigachat/login>

из-за углеродного следа, возникающего вследствие потребления современным оборудованием огромного количества энергии, необходимой для тензорной обработки.

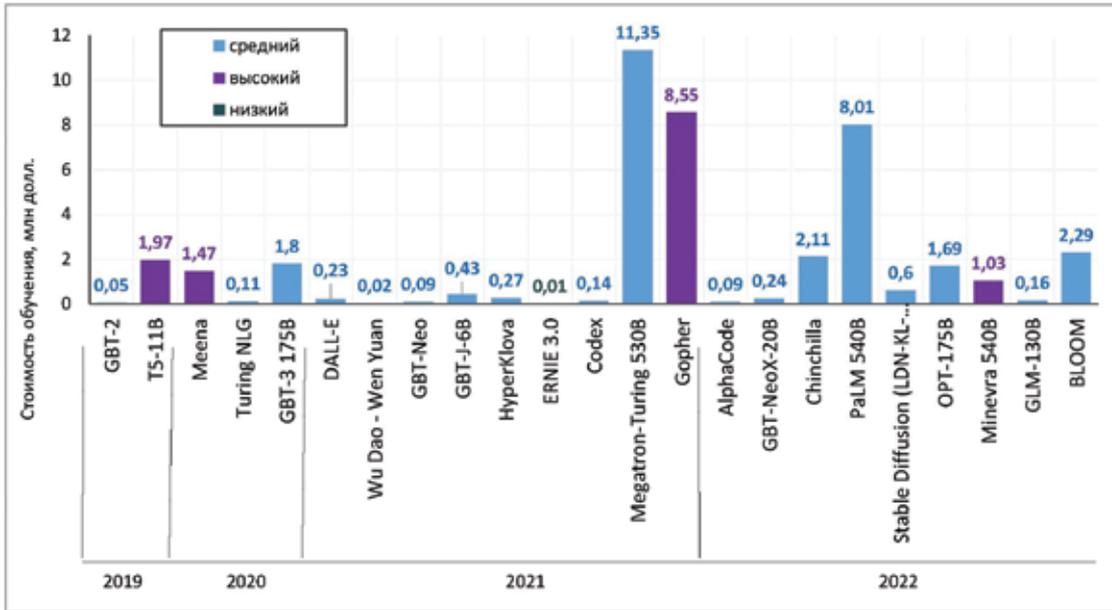
Основным вопросом, связанным с современными крупноразмерными системами искусственного интеллекта, является их высокая стоимость. Компании-разработчики неохотно раскрывают свои затраты, однако по информации о стоимости машинного обучения уже можно сделать вывод, что создание систем искусственного интеллекта ушло из сферы индивидуального или группового «подвижного» творчества и стало объектом коммерческого интереса крупных институциональных участников рынка, способных обеспечить инженерный процесс необходимыми финансовыми и технологическими ресурсами. Так, машинное обучение модели BLOOM составило 2,3 млн долл., а крупноразмерной языковой системы Chinchilla, созданной в мае 2022 г. компанией DeepMind – 2,1 млн долл. (рис. 1). 2023 г. принес новые рекорды: обучение модели GPT-4 (компания OpenAI) потребовало вычислительных ресурсов стоимостью в 78,3 млн долл., а обучение модели Gemini Ultra обошлось создателю (компания Google) в 191,4 млн долл.⁸

Второй вопрос – ресурсоемкость современных интеллектуальных систем. Активное распространение, а также увеличение размерности систем искусственного интеллекта формируют необходимость осознанного отслеживания их ресурсопотребления и влияния на окружающую среду. Оптимизация расхода ресурсов сама по себе является важным направлением применения искусственного интеллекта, однако его использование дает мультипликативный эффект, направленный на повышение эффективности самих систем искусственного интеллекта, где машинное обучение и вывод обеспечиваются огромными вычислительными процессами, требующими соответствующих энергозатрат. Так, обучение системы BLOOM потребовало затрат энергии, сопоставимых с потреблением среднего американского домохозяйства в течение 41 года⁹.

Компания DeepMind в 2021 г. провела эксперимент, в ходе которого специально созданная программа BCOOLER (BVE-based Constrained Optimization Learner with Ensemble Regularization), рассчитанная на машинное обучение с подкреплением, была использована для оптимизации процессов охлаждения центров обработки данных компании Google. По итогам 3-месячного

⁸The AI Index 2024 Annual Report // AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA: April 2024. URL: https://aiindex.stanford.edu/wp-content/uploads/2024/04/HAI_AI-Index-Report-2024.pdf (дата обращения: 18.04.2024).

⁹Прим. Автора: По оценке Управления энергетической информации США, среднегодовое потребление электроэнергии американскими потребителями коммунальных услуг составляет 10,632 кВт-ч. См.: Residential energy consumption survey // U.S. Energy Information Administration. URL: <https://www.eia.gov/consumption/residential/data> (дата обращения: 15.01.2024); *The AI index 2023 annual report* // AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA. April 2023. URL: <https://hai.stanford.edu/research/ai-index-2023>. URL: <https://hai.stanford.edu/research/ai-index-2023> (дата обращения: 10.02.2024).



Примечание:

Исследователи AI Index каждому показателю назначали рейтинг, исходя из степени экспертного доверия к достоверности данных о стоимости машинного обучения: «средний» (оценка, близкая к реальной), «низкий» (оценка занижена), «высокий» (оценка завышена). В случаях нехватки данных модели исключались из анализа.

Источник: *The AI index 2023 annual report // AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA. April 2023. URL: <https://hai.stanford.edu/research/ai-index-2023> (дата обращения: 10.02.2024).*

Рис. 1. Стоимость машинного обучения крупноразмерных мультимодальных моделей, млн долл. США

Source: *The AI index 2023 annual report. AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA. April 2023. URL: <https://hai.stanford.edu/research/ai-index-2023> (accessed: 10.02.2024)*

Fig. 1. Cost of machine learning of large-scale multimodal models, USD million

обучения программа достигла экономии энергии в 12,7%, не снизив при этом комфортного уровня охлаждения [16]. Таким образом, сэкономленная электроэнергия в пересчете на потребление системы BLOOM составила объем, способный обеспечить потребление среднего домохозяйства в США в течение более чем 5-ти лет.

Вышеуказанные результаты имеют не только экономическое, но и экологическое измерение. В настоящее время не существует общепринятого стандарта обозначения величины «углеродного следа» от систем искусственного интеллекта, и для сравнения исследователи в качестве базовых характеристик используют углеродоемкость системы, энергозатраты и производные от них показатели. Одним из таких показателей является «эффективность потребления энергии» (PUE, Power Usage Efficiency), где рост числового значения означает снижение энергоэффективности центра обработки данных. PUE рассчитывается как отношение общего энергопотребления дата-центра (включая дополнительную энергию, необходимую для поддержки вычислительной инфраструктуры, прежде всего, для охлаждения системы и кондиционирования воздуха) к объему энергии, необходи-

мой непосредственно для работы компьютерного оборудования:

$$PUE = \frac{p_1 + p_2 + \dots + p_n}{p_1} \quad (2)$$

где p_1 – энергия, потребляемая компьютерным оборудованием; p_2, \dots, p_n – дополнительная энергия, необходимая для обеспечения работы вычислительной инфраструктуры.

Сравнительный анализ, представленный в табл. 2, дает возможность косвенно оценить «углеродный след» 4-х крупных интеллектуальных систем: BLOOM, OPT, Gopher, GPT-3. «Лидером» в данной группе выступает модель GPT-3, потребляющая углерода в 20,1 раз больше, чем BLOOM, и в 7,2 и 1,4 раза больше, чем OPT и Gopher соответственно [24]. Самое «безобидное» из указанных построений – BLOOM – в ходе обучения выбросило в атмосферу углекислого газа в 25 раз больше, чем диоксида углерода вырабатывается за один авиационный перелет из Нью-Йорка в Сан-Франциско (в обе стороны) из расчета на одного пассажира, или в 1,4 раза больше «углеродного следа» жителя США за 12 месяцев (рис. 2) [25].

Таблица 2

Оценка техногенного воздействия на окружающую среду в результате машинного обучения различных моделей искусственного интеллекта

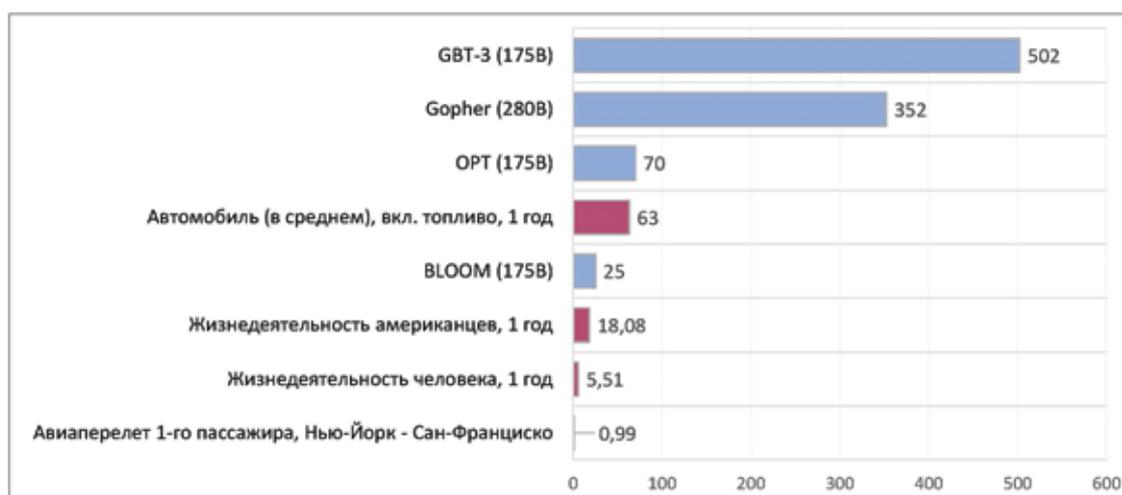
Table 2

Estimated anthropogenic environmental impact from machine learning of different artificial intelligence models

Модель	Число параметров	PUE	Углеродоемкость сети	Потребляемая энергия	Выбросы в эквиваленте	
					CO ₂	CO ₂ x PUE
Gopher	280B	1.08	330 gCO ₂ eq/kWh	1,066 MWh	352 тонн	380 тонн
BLOOM	176B	1.20	57 gCO ₂ eq/kWh	433 MWh	25 тонн	30 тонн
GPT-3	175B	1.10	429 gCO ₂ eq/kWh	1,287 MWh	502 тонн	552 тонн
OPT	175B	1.09	231 gCO ₂ eq/kWh	324 MWh	70 тонн	76.3 тонн

Источник: [24].

Source: [24].



Источник: [25].

 Рис. 2. Выбросы CO₂ в окружающую среду в результате машинного обучения различных моделей искусственного интеллекта, тонны

Source: [25].

 Fig. 2. CO₂ emissions from machine learning of different artificial intelligence models, tonnes

Основная трудность применения «традиционных», более ранних интеллектуальных систем заключается в том, что, при вполне корректном решении узкопрофильных вопросов, они испытывали большие сложности при возникновении более широких задач. Новые же модели (как, например, Gato, PaLi, BEiT-3) выступают как интеллектуальные системы, способные работать одновременно над несколькими задачами (например, язык, зрение). Тем не менее, они также испытывают большое количество проблем. В отчете Стэнфордского университета о развитии искусственного интеллекта

за 2023 г.¹⁰ перечислены нижеследующие технологические проблемы функционирования новейших интеллектуальных построений.

1. Системы искусственного интеллекта продолжают работать над самыми современными задачами, но прогресс в изменении многих показателей остается незначительным в сравнении с показателями предыдущего года.
2. Новейшие чат-боты (ChatGPT), модели преобразования «текст-видео» (Make-A-Video) и «текст-изображение» (Stable Diffusion, DALL-E 3)

¹⁰The AI index 2023 annual report // AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA. April 2023. URL: <https://hai.stanford.edu/research/ai-index-2023> (дата обращения: 10.02.2024).

иногда допускают галлюцинации, выдают сложные или бессвязные ответы. Указанное обстоятельство затрудняет их применение в критически важных задачах.

3. Генеративные способности языковых систем продолжают совершенствоваться, однако модели все еще испытывают значительные трудности при возникновении сложных задач, требующих рассуждения.
4. Также генеративные модели испытывают трудности при решении сложных задач планирования.

Экономические аспекты искусственного интеллекта

Расширение технологических возможностей систем искусственного интеллекта повлекло наращивание темпов внедрения указанных технологий в коммерческой и общественной жизни, а также в сфере государственного управления. Столь активное расширение спроса способствовало огромному притоку частных инвестиций в сектор интеллектуальных технологий, объем которых всего за десятилетие (2013–2022 гг.) вырос в 18 раз¹¹. Несмотря на то, что после почти двукратного увеличения данного показателя в 2021 г. (до 132,36 млрд долл.) в последующий период наблюдалось некоторое охлаждение (точнее, коррекция) частной инвестиционной активности (103,36 млрд долл. в 2022 г. и 95,99 млрд долл. в 2023 г), совокупный объем частных вложений в искусственный интеллект в настоящий момент почти на 50% превышает уровень в 64,02 млрд долл., достигнутый в 2020 г., предшествовавшем резкому взлету инвестиционной активности 2021 г. Примечательно, что на фоне общего снижения объема частных инвестиций в 2023 г. резко увеличилось (почти в 9 раз по сравнению с 2022 г. и в 30 раз – с 2019 г.) финансирование генеративного искусственного интеллекта (25,2 млрд долл.), на который в 2023 г. пришлось более 25% всех частных инвестиций, связанных с искусственным интеллектом.

Лидером в общем объеме частных вложений в технологии искусственного интеллекта остаются США, которые в 2023 г. инвестировали 67,22 млрд долл., в 8,6 раза превысив показатель Китая, занимающего второе место, с объемом инвестиций в 7,76 млрд долл. (в 2022 г. в 3,5 раза боль-

ше – 47,4 и 13,4 млрд долл. соответственно). На третьем месте находится Англия с объемом инвестиций в 3,78 млрд долл. (в 17,8 раз меньше США). Если учитывать общую сумму инвестиций за десятилетие, с 2013 г., то состав лидеров не изменится: 335,24 млрд долл. – США, 103,65 млрд – Китай, 22,25 млрд – Англия (рис. 3).

В 2023 г. больше всего частных инвестиций поступило по направлениям: «инфраструктура, исследование и управление искусственным интеллектом» (18,3 млрд долл.); «обработка естественного языка и поддержка клиентов» (8,1 млрд долл.); «управление и обработка данных» (5,5 млрд долл.). Значительная доля первого направления отражает крупные инвестиции в компании, специализирующиеся на разработке приложений для систем искусственного интеллекта (например, Anthropic, Inflection AI, OpenAI).

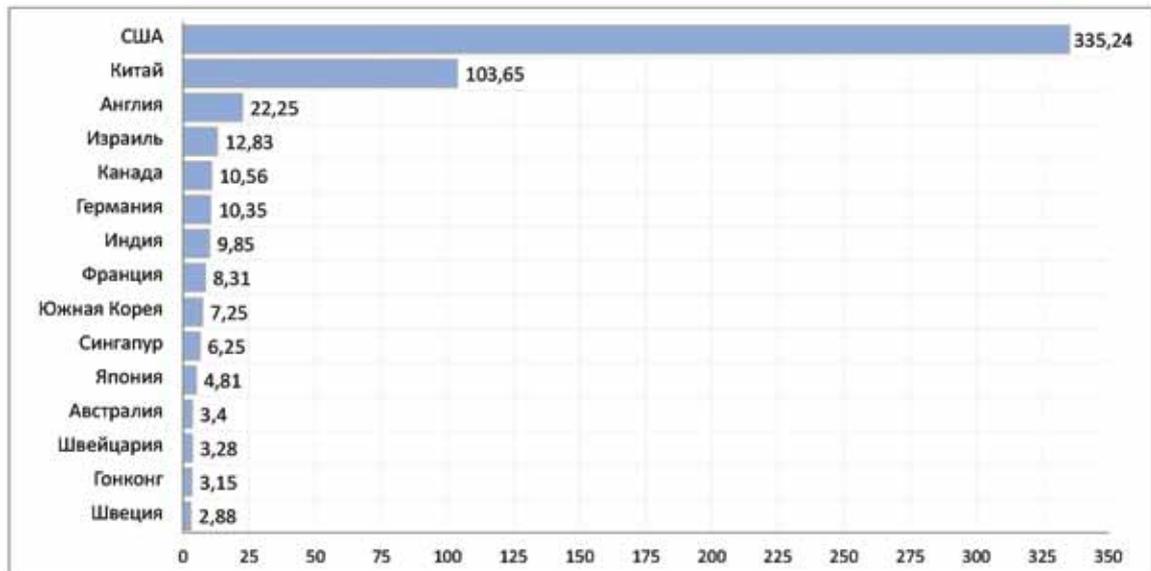
США заметно опережали остальные страны почти на всех направлениях частного инвестирования в искусственный интеллект. Исключением стало направление «распознавание лиц», где американские вложения в 2023 г. уступали китайским (0,09 млрд / 0,13 млрд долл.). Сопоставимые инвестиции оба технологических лидера имели по направлению «полупроводники»: США – 0,79 млрд, Китай – 0,63 млрд долл. При этом в предыдущем году инвестиции Китая в полупроводники составили 1,02 млрд долл., что в 1,75 раза превысило американские и в 102 раза совокупные европейские и британские вложения (0,58 млрд и 0,01 млрд долл.).

Несмотря на американское и китайское лидерство в объемах частных инвестиций в создание систем искусственного интеллекта, в 2022 г. отмечалась отрицательная динамика обеих стран в этом показателе: -35,5% в США и -41,3% в Китае (по сравнению с предыдущим годом)¹². В 2023 г. показатель Китая продолжил падение (-44,2%), но США увеличили годовые частные инвестиции в данный сектор на 22,1%¹³. Источниками столь существенного снижения данного показателя могут выступать различные причины как экономического, так и неэкономического свойства (например, фаза спада инвестиционного цикла или возрастающая внешнеполитическая неопределенность). Однако можно однозначно утверждать, что 2-летний период может не обладать необходимой репрезентативностью, а на длительном периоде частные

¹¹ The AI index 2023 annual report // AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA. April 2023. URL: <https://hai.stanford.edu/research/ai-index-2023> (дата обращения: 10.02.2024).

¹² The AI index 2023 annual report // AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA. April 2023. URL: <https://hai.stanford.edu/research/ai-index-2023> (дата обращения: 10.02.2024).

¹³ The AI Index 2024 Annual Report // AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA: April 2024. URL: https://aiindex.stanford.edu/wp-content/uploads/2024/04/HAI_AI-Index-Report-2024.pdf (дата обращения: 18.04.2024).



Источник: *The AI Index 2024 Annual Report* // AI Index Steering Committee, Institute for Human-Centered AI, Stanford University. Stanford, CA: April 2024. URL: https://aiindex.stanford.edu/wp-content/uploads/2024/04/HAI_AI-Index-Report-2024.pdf (дата обращения: 18.04.2024).

Рис. 3. Частные инвестиции в разработку искусственного интеллекта по странам, млрд долл. США, всего за 2013–2023 г.

Source: *The AI Index 2024 Annual Report* // AI Index Steering Committee, Institute for Human-Centered AI, Stanford University. Stanford, CA: April 2024. URL: https://aiindex.stanford.edu/wp-content/uploads/2024/04/HAI_AI-Index-Report-2024.pdf (accessed: 18.04.2024)

Fig. 3. Private investments in the development of artificial intelligence by countries, USD billion, total for 2013–2023

инвестиции в искусственный интеллект сохраняют устойчивую положительную динамику. Кроме того, интерес инвесторов к технологиям искусственного интеллекта подтверждается ростом финансовых вложений в новые ИТ-компании и ИТ-стартапы.

По совокупному числу новых компаний (стартапов), созданных в 2013–2023 гг. в секторе искусственного интеллекта, лидируют США. Их количество в 3,8 раза превышает число таких компаний в Китае (5509 / 1446 компаний) и почти в 7,6 раза – в Англии (5509 / 727 компаний). По данному показателю только в 2023 г. Китай отставал от США более чем в 7,3 раза (122 / 897 компаний), Великобритания – в 8,8 раза (104 / 897 компаний).

В 2023 г. глобальный объем инвестиционных вложений в новые проекты в области искусственного интеллекта составил 42,5 млрд долл., что на 10% ниже аналогичного показателя предыдущего года¹⁴. Указанный факт не выглядит критично на фоне 38%-го сокращения объема инвестиций

в стартапы по всем направлениям (в 2022–2023 гг.: 462 млрд. / 285 млрд. долл.). Примерно 2/3 указанных инвестиций поступили от корпорации Microsoft (1,3 млрд долл. – в компанию Inflection AI; 10 млрд – разработчику системы ChatGPT компании OpenAI), а также от Amazon и Google (6 млрд долл.), которые синдицированно профинансировали разработки компании Anthropic, занимающейся построением больших языковых моделей¹⁵. Указанное обстоятельство демонстрирует тот факт, что крупные технологические компании обошли в 2023 г. венчурные фонды по объемам финансовых вложений в новые разработки искусственного интеллекта. Впрочем, молодые компании-разработчики охотно аффилируются в системы крупных корпораций, так как последние, помимо инвестиций, могут обеспечить дополнительные технологические возможности (облачную инфраструктуру, аппаратно-вычислительные мощности и проч.). Основатель венчурного фонда Tapestry VC П. Мерфи признается, что «на фоне

¹⁴ State of AI 2023 Report // CB Insights. URL: <https://www.cbinsights.com/research/report/ai-trends-2023> (дата обращения: 12.03.2024).

¹⁵ Teare G. Global Startup Funding in 2023 Clocks in at Lowest Level in 5 Years // Crunchbase News. URL: <https://news.crunchbase.com/venture/global-funding-data-analysis-ai-eoy-2023> (дата обращения: 10.02.2024).

стремительного развития платформ искусственного интеллекта технологические гиганты взяли под контроль большинство стартапов, обладающих исключительным потенциалом. Ведущие мировые венчурные фонды, управляющие десятками миллиардов долларов, не в состоянии бороться за сохранение независимости этих компаний»¹⁶.

Следует указать на еще одно немаловажное изменение на рынке искусственного интеллекта. До 2014 г. наиболее заметные системы создавались академическими учреждениями. Однако современные модели становятся все крупнее и затратнее. Так, построенная в 2022 г. флагманская модель Google PaLM по сравнению с одной из первых аналогичных моделей GPT-2 (2019 г.) оказалась в 360 раз больше и в 160 раз дороже. Указанные трансформации требуют огромных дополнительных финансовых, вычислительных и информационных ресурсов, которыми научные и некоммерческие организации нередко не располагают. Данное обстоятельство привело к переформатированию рынка искусственного интеллекта, которое можно проиллюстрировать тем фактом, что в 2022 г. коммерческие ИТ-компании представили 32, а академические круги – лишь 3 заметные интеллектуальные системы¹⁷. В следующем, 2023 г., разрыв сократился, но оставался существенным – 51 и 15 соответственно. Одновременно промышленность и научная сфера в кооперации построили дополнительно 21 подобную модель¹⁸, что тоже является достойной внимания тенденцией.

В 2024 г. объем глобального рынка искусственного интеллекта составил 214,6 млрд долл. К 2030 г. он увеличится, согласно экспертным оценкам, до 1339,1 млрд долл., что продемонстрирует среднегодовой темп роста в 35,7%¹⁹.

На этом фоне Россия не выглядит впечатляюще. Согласно оценке проектного офиса по реализа-

ции программы «Цифровая экономика Российской Федерации»²⁰, отечественный рынок по итогам 2023 г. составил 650 млрд руб.²¹, или примерно 7,2 млрд долл., что равно доле мирового рынка искусственного интеллекта в 3,36%. При всей неамбициозности указанного показателя к нему следует относиться настороженно, так как элементарная экстраполяция на Россию среднегодового темпа роста мирового рынка искусственного интеллекта за период 2019–2023 гг. (29,8%) оставляет ей долю лишь в 0,5% (как и в 2019 г., с показателями объемов российского и мирового рынка искусственного интеллекта в 0,291 и 58,3 млрд долл.). Возникает закономерный вопрос, какие достижения обеспечили отечественному сектору искусственного интеллекта среднегодовой темп роста, в 3 раза превышающий общемировой? Возможно, экспертам программы «Цифровая экономика» и госслужащим, отвечающим за развитие ИТ-сферы, известны какие-то подробности, но у непосвященного исследователя заявленные цифры вызывают обоснованное недоверие.

Россия в своих разработках искусственного интеллекта заметно отстает от ведущих стран, что признается и в отечественном экспертном сообществе. Согласно так называемому «Индексу зрелости технологий искусственного интеллекта», разработанному рабочей группой Федерального проекта по искусственному интеллекту на основе показателей ИТ-инфраструктуры, уровня развития науки, обеспеченности научно-техническими кадрами в стране, доходов национальных ИТ-компаний и проч., Россия имеет лишь 6 баллов (США и Китай – по 77, Великобритания, Канада, Германия – по 26 баллов).

Несмотря на отдельные модели, о которых упоминалось выше, российский вклад в глобальное развитие искусственного интеллекта не признается внешними наблюдателями в качестве определя-

¹⁶ Big tech outspends venture capital firms in AI investment frenzy // Financial Times. 29.12.2023. URL: <https://www.ft.com/content/c6b47d24-b435-4f41-b197-2d826cce9532> (дата обращения: 10.02.2024).

¹⁷ The AI index 2023 annual report // AI Index Steering Committee, Institute for Human-Centered AI, Stanford University, Stanford, CA. April 2023. URL: <https://hai.stanford.edu/research/ai-index-2023> (дата обращения: 10.02.2024).

¹⁸ The AI Index 2024 Annual Report // AI Index Steering Committee, Institute for Human-Centered AI, Stanford University. Stanford, CA: April 2024. URL: https://aiindex.stanford.edu/wp-content/uploads/2024/04/HAI_AI-Index-Report-2024.pdf (дата обращения: 18.04.2024).

¹⁹ Artificial Intelligence (AI) Market by Offering (Discriminative AI, Generative AI, Hardware, Services), Technology (ML, NLP, Context-aware AI, Computer Vision), Business Function (Marketing & HR), Vertical and Region - Global Forecast to 2030 // MarketsandMarkets. URL: <https://www.marketsandmarkets.com/Market-Reports/artificial-intelligence-market-74851580.html> (дата обращения: 31.05.2024).

²⁰ Прим. Автора: Согласно постановлению Правительства РФ «О системе управления реализацией национальной программы «Цифровая экономика Российской Федерации» № 234 от 02.03.2019 г., функции проектного офиса по реализации национальной программы «Цифровая экономика Российской Федерации» возложены на автономную некоммерческую организацию «Аналитический центр при Правительстве Российской Федерации».

²¹ Искусственный интеллект (рынок России) // TAdviser. URL: [https://www.tadviser.ru/index.php/Статья:_Искусственный_интеллект_\(рынок_России\)](https://www.tadviser.ru/index.php/Статья:_Искусственный_интеллект_(рынок_России)) (дата обращения: 31.05.2024).

шего. Соответствующие ежегодные исследования, например, Стэнфордского университета или компании McKinsey, не упоминают достижения российских разработчиков. В многостраничном обзоре компании MarketsandMarkets²² описаны успехи 65-ти компаний (46 – из США, по 3 – Китай, ФРГ и Израиль, по 2 – Англия и Канада, и т.д.). Ни одна из российских компаний в фокус исследования не попала, что может показаться несправедливым с учетом имеющейся информации об активности российских компаний в разработках искусственного интеллекта. Вместе с тем, представляется, что данное обстоятельство не следует воспринимать эмоционально, в качестве желания зарубежных наблюдателей принизить российские достижения, а нужно принять как реальный факт, требующий глубокого осмысления, возможно, в ходе отдельного исследования, посвященного развитию искусственного интеллекта в России.

Этические аспекты искусственного интеллекта

Как зачастую происходит в мире нововведений, внедрение искусственного интеллекта наряду с получаемой пользой формирует и обратный эффект.

Согласно данным Репозитория AIAAIC²³, число сообщений о проблемах, связанных с использованием данных технологий, лишь за одно десятилетие, с 2012 г., увеличилось в 26 раз. Столь заметный рост данного показателя может свидетельствовать как о совершенствовании системы учета указанных инцидентов, так и о возрастающей степени вовлеченности технологий искусственного интеллекта в общественную жизнь. Однако в любом случае подобный рост демонстрирует обострение этических проблем, связанных с развитием искусственного интеллекта. В 2023 г. AIAAIC зарегистрировал 123 инцидента, что на 32,3 % больше, чем в предыдущем году.

По данным компании McKinsey²⁴, представители бизнеса рассматривают как наиболее подверженную рискам при внедрении искусственного интеллекта сферу кибербезопасности (59%). Далее следуют: нормативно-правовое соответствие

(45%), конфиденциальность личности (40%), объяснимость процессов (37%), репутация организации (32%), равноправие и справедливость (30%), вытеснение рабочей силы (28%). Замыкают «линейку» озабоченности бизнеса политическая стабильность (9%) и национальная безопасность (13%). Кибербезопасность, нормативно-правовое соответствие, конфиденциальность личности лидируют также и среди направлений, по которым представители бизнеса предпринимают практические усилия по снижению рисков (51%, 36% и 28% соответственно).

Многие эксперты, говоря о рисках, порождаемых искусственным интеллектом, указывают на его влияние на рынок труда. Возникновение опасений о своем будущем в преддверии технологических изменений характерно для всех поколений. «Неолуддиты» в наше время также заявляют о будущем катастрофическом падении занятости вследствие развития систем искусственного интеллекта.

Действительно, если предполагаемый потенциал генеративного искусственного интеллекта окажется реализуемым, рынок труда может подвергнуться значительным потрясениям. Эксперты банка Goldman Sachs утверждают, что автоматизации в различной степени будут подвержены 2/3 существующих рабочих мест, а сами системы генеративного искусственного интеллекта способны заместить до 1/4 объема выполняемой в настоящее время работы. Экстраполяция указанных оценок на глобальный уровень дает исследователям основания заявить о возможной автоматизации производственных процессов, обеспечивающих в настоящее время 300 млн рабочих мест с полной занятостью²⁵.

Вместе с тем, установлено, что технологические инновации в долгосрочном плане обеспечивают рост занятости, а текущее сокращение рабочих мест в результате автоматизации компенсируется появлением новых профессий и соответствующим созданием новых рабочих мест. Так, Аутор Д. и др. на основе обработки данных переписи насе-

²² Artificial Intelligence (AI) Market by Offering (Discriminative AI, Generative AI, Hardware, Services), Technology (ML, NLP, Context-aware AI, Computer Vision), Business Function (Marketing & HR), Vertical and Region – Global Forecast to 2030 // MarketsandMarkets. URL: <https://www.marketsandmarkets.com/Market-Reports/artificial-intelligence-market-74851580.html> (дата обращения: 31.05.2024).

²³ Прим. Автора: Репозиторий инцидентов и противоречий в области искусственного интеллекта, алгоритмов и автоматизации (AIAAIC – The AI, Algorithmic, and Automation Incidents and Controversies Repository) был запущен как частный проект в 2019 г. Он представляет собой открытую и независимую базу данных об этических проблемах, возникающих в ходе эксплуатации соответствующих технологий. (См. AIAAIC Repository. URL: <https://www.aiaaic.org/aiaaic-repository>).

²⁴ The state of AI in 2022 – and a half decade in review // QuantumBlack by McKinsey. URL: <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review> (дата обращения: 23.02.2024).

²⁵ Briggs J., Kodnani D. Global economics analyst the potentially large effects of artificial intelligence on economic growth // Goldman Sachs. 26.03.2023. URL: https://www.key4biz.it/wp-content/uploads/2023/03/Global-Economics-Analyst_-The-Potentially-Large-Effects-of-Artificial-Intelligence-on-Economic-Growth-Briggs_Kodnani.pdf (дата обращения: 15.03.2024).

ления установили, что 60% работающих граждан в настоящее время трудится в профессиях, не существовавших в 1940 г., а развитие технологий в 1940–2018 гг. обеспечило 85% расширения рынка труда [15].

Согласно исследованию Асемоглу Д. и Рестрепо П. [26], внедрение новых технологий в 1950–1970-е гг.

обеспечивало рост производительности труда, который происходил на фоне вытеснения работников из «обновляемых» секторов экономики, однако одновременно аналогичными темпами создавались новые рабочие места. С 1980-х гг. последний процесс затормозился, и сокращение рабочих мест вследствие технологических трансформаций опережало создание новых рабочих мест (рис. 4).



Источник: [26].

Рис. 4. Вклад новых технологий в общий рост фонда заработной платы

Source: [26].

Fig. 4. Contribution of new technologies to the total growth of the wage fund

Результаты проведенных исследований дают основания полагать, что развитие систем генеративного искусственного интеллекта способно в ближайшей перспективе оказать на рынок труда отрицательный эффект (по аналогии с внедрением более ранних информационных технологий), что, тем не менее, будет происходить на фоне общего повышения производительности труда. По расчетам экспертов Goldman Sachs, производительность труда США в результате широкого внедрения систем генеративного искусственного интеллекта может в течение 10-ти лет продемонстрировать рост в пределах около 1,5 процентных пункта²⁶.

Вместе с тем, существуют более серьезные риски, связанные с бесконтрольным развитием искусственного интеллекта, которые пока имеют гипотетический характер, но их потенциальная опасность требует осмысления уже в настоящий момент.

Научно-технический прогресс во все времена так или иначе был связан с попытками воспроизведения функций человеческого организма вспомогательными средствами. При этом творческая энергия разработчиков была ориентирована на максимальное замещение мышечных усилий человека в процессе осуществления какой-либо физической деятельности. Технологические нововведения выступали в данном случае в качестве вспомогательного инструментария, не ставящего под сомнение приоритет человека при выработке и принятии управленческих решений.

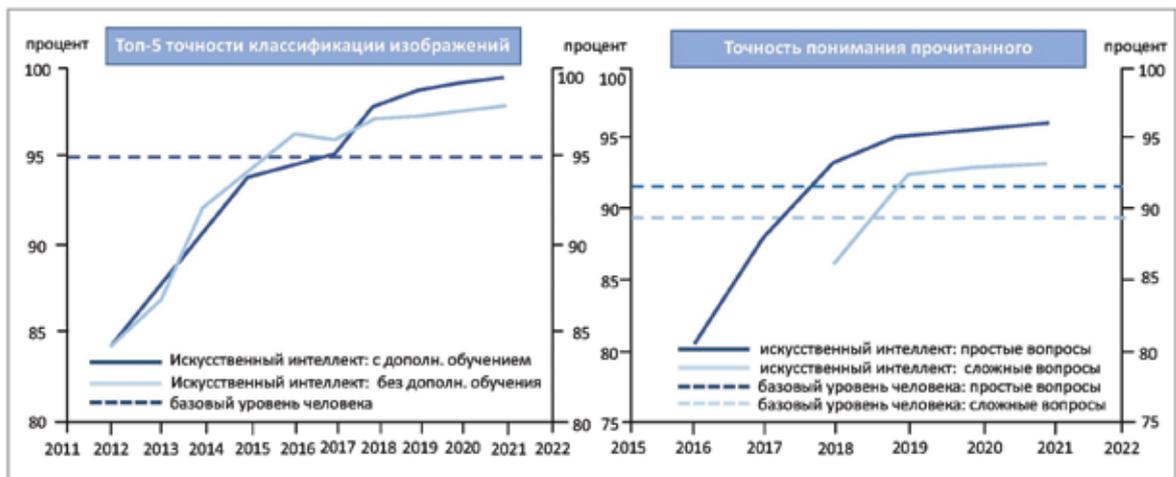
Экспоненциальное развитие технологий искусственного интеллекта, происходящее в настоящий период технологического развития общества, ориентировано на максимальное воспроизводство когнитивных способностей человека и постепенную передачу «умным машинам» ключевых компетенций в области

²⁶ Briggs J., Kodnani D. Global economics analyst the potentially large effects of artificial intelligence on economic growth // Goldman Sachs. 26.03.2023. URL: https://www.key4biz.it/wp-content/uploads/2023/03/Global-Economics-Analyst-The-Potentially-Large-Effects-of-Artificial-Intelligence-on-Economic-Growth-Briggs_Kodnani.pdf (дата обращения: 15.03.2024).

принятия решений, что способно генерировать риски утраты человеком субъектности по отношению к окружающим процессам и превращению его в объект приложения когнитивной деятельности создаваемых интеллектуальных систем.

Предпосылки реализации столь тревожных перспектив уже имеются. Так, алгоритмы, заложенные

в основу систем генеративного искусственного интеллекта, начали превосходить собственные показатели человека в реализации таких когнитивных задач как понимание содержания текстов, классификация изображений и прочее. Указанный факт наглядно иллюстрирует информация, представленная на рис. 5.



Источник: Briggs J., Kodnani D. Global economics analyst the potentially large effects of artificial intelligence on economic growth // Goldman Sachs. 26.03.2023. URL: https://www.key4biz.it/wp-content/uploads/2023/03/Global-Economics-Analyst_-The-Potentially-Large-Effects-of-Artificial-Intelligence-on-Economic-Growth-Briggs_Kodnani.pdf (дата обращения: 29.03.2024).

Рис. 5. Сравнение возможностей искусственного интеллекта с базовым уровнем человека

Source: Briggs J., Kodnani D. Global economics analyst the potentially large effects of artificial intelligence on economic growth. Goldman Sachs. 26.03.2023. URL: https://www.key4biz.it/wp-content/uploads/2023/03/Global-Economics-Analyst_-The-Potentially-Large-Effects-of-Artificial-Intelligence-on-Economic-Growth-Briggs_Kodnani.pdf (accessed: 29.03.2024)

Fig. 5. Comparison of artificial intelligence capabilities with the basic human level

Технологии искусственного интеллекта становятся все сложнее и приобретают все большую автономность от воли человека. Данное обстоятельство выдвигает на передний план вопросы соблюдения человеческой морали и этики при принятии интеллектуальными системами управленческих решений.

Однако при перенесении на машины норм человеческой морали следует учитывать важное обстоятельство. Действия искусственного интеллекта представляют собой, по сути, реализацию жестких математических алгоритмов, где ограничения, воспроизводящие нормы морали и человеческой этики, должны быть также жестко прописаны. Однако сама человеческая мораль не имеет четко прописанных правил, границы ее реализации достаточно лабильны и в схожих ситуациях могут существенно изменяться в зависимости от социокультурных установок конкретного индивида. В этом видится основная сложность перенесения на искусственный интеллект тех установок, которые сам человек еще не определил для себя в доста-

точно различимой, институционально оформленной, признаваемой всеми участниками общества конструкции морально-этических норм.

Более того, даже в случае институционализации этических норм, их разработка, вероятно, не станет простой экстраполяцией на машины классической этики, применяемой в отношениях между людьми. Результаты экспериментов демонстрируют, что люди в вопросе соблюдения этических норм предъявляют в отношении «думающих машин» более весомые ожидания, чем те стандарты поведения, которыми в аналогичных обстоятельствах руководствуется сам человек [27, с. 122].

С технологической точки зрения формальные вопросы этики должны находиться в компетенции разработчиков. Однако последние могут не обладать необходимой мотивацией для учета этических вопросов при разработке соответствующей технологии. Более того, до сих пор учет этических вопросов, обусловленных возможными рисками внедрения новой технологии, не является обязательной составляющей комплекса работ, осу-

ществляемых инженерами и управленцами. Более того, они искренне могут не осознавать этических проблем, способных возникнуть по итогам реализации перспективной разработки.

Наверное, первым заметным свидетельством озабоченности экспертного, научного и делового сообществ нарастающей проблемой этики в вопросах создания систем искусственного интеллекта стало открытое письмо Института будущего человечества (Future of Humanity Institute), опубликованное в марте 2023 г. Свою подпись под данным обращением поставили Стивен Возняк (соучредитель Apple), Илон Маск (генеральный директор SpaceX и Tesla, основатель xAI), Эмад Мостак (генеральный директор Stability AI) и еще более тысячи представителей сферы интеллектуальных технологий. Подписавшие обращение призвали крупных разработчиков искусственного интеллекта приостановить обучение систем, «более мощных, чем GPT-4»²⁷, пока не будет определен вопрос контроля за развитием таких построений.

Мировой опыт конструкторской мысли содержит примеры добровольных самоограничений при разработке новых технологий. Так, в этическом кодексе Национального общества профессиональных инженеров (США) первым «фундаментальным каноном» является то, что инженеры должны «ставить во главу угла безопасность, здоровье и благосостояние общества»²⁸. Если введение моральных норм для машин повысит уровень общественного благосостояния и безопасности, то американские инженеры, согласно их собственному этическому кодексу, обязаны это сделать.

Понятно, что положение вышеуказанного этического кодекса не ориентировано на простейшие приспособления, которые не обладают ни самостоятельностью, ни чувствительностью к ценностям. Их применение полностью обусловлено волей и целями человека, использующего тот или иной инструмент. Но уже даже самые простые технологии могут содержать в своей конструкции некую «операционную мораль», предусмотренную разработчиком. Например, предохранитель в боевом или охотничьем оружии не влияет на работоспособность механизма, однако ориентирован на недопущение случайного или самопроизвольного выстрела. В данном случае внедрение в технологию операционной морали полностью подчинено

разработчику, а сам процесс проектирования технологии произведен с непосредственным учетом этических норм.

На другой стороне процесса развития современных технологий находятся полностью автономные, обладающие высокой чувствительностью к ценностям интеллектуальные системы, способные выступать в окружающем социальном пространстве как надежные (ответственные) моральные агенты. Данные системы, обладающие «полной ответственной моралью» [17], способны принимать решения и действовать как человек, на основе заложенных этических установок. Возможность создания подобных систем является гипотетическим вопросом, однако логика научно-технического прогресса нацелена именно на этот результат.

По мере распространения технологий автономного принятия решений резонно ожидать от разработчиков, что создание этических ограничений в искусственном интеллекте станет инженерным императивом.

Также можно наблюдать явные признаки озабоченности рассматриваемой проблемой со стороны государственных властей. Например, в США – стране, лидирующей по созданию интеллектуальных систем – количество документов, регулирующих отношения в сфере искусственного интеллекта, заметно увеличилось и составляет сегодня 25 нормативных актов, по сравнению с одним в 2016 г. Только в 2023 г. совокупное количество регуляторных актов, связанных с искусственным интеллектом, выросло на 56,3%²⁹.

Однако сложно оценить, насколько устойчивыми на практике окажутся принципы государственных властей и самоограничения инженеров перед творческими амбициями разработчиков и вопросами рыночной окупаемости новой технологии. А если речь пойдет о каких-либо аспектах межгосударственной конкуренции в разработке того или иного инновационного направления, особенно в оборонной сфере, можно предположить, что вопросы этики создания новой технологии будут «отодвинуты» на последний план.

Выводы

Проведенный анализ демонстрирует двойственный характер внедрения систем искусственного

²⁷ Pause Giant AI Experiments: An Open Letter // Future of Life Institute. 22.03.2023. URL: <https://www.futureoflife.org/open-letter/pause-giant-ai-experiments> (дата обращения: 31.03.2023)

²⁸ NSPE Code of Ethics for Engineers // National Society of Professional Engineers. URL: <https://www.nspe.org/resources/ethics/code-ethics> (дата обращения: 15.12.2023).

²⁹ The AI Index 2024 Annual Report // AI Index Steering Committee, Institute for Human-Centered AI, Stanford University. Stanford, CA: April 2024. URL: https://aiindex.stanford.edu/wp-content/uploads/2024/04/HAI_AI-Index-Report-2024.pdf (дата обращения: 18.04.2024).

интеллекта, который находит убедительное отражение как в экономическом или технико-технологическом, так и в этическом измерениях.

С одной стороны, разработка и внедрение указанных систем открывает широкие перспективы применения «умных» машин в общественной жизни. Однако, при всей неоспоримости прогрессивного влияния указанных технологий на общественное развитие, следует учитывать ряд угроз и рисков, имеющих в настоящее время как гипотетический, так и уже вполне очевидный характер.

Так, в экономической плоскости указанная дихотомия прослеживается в негативной тенденции к олигополизации отношений в секторе высокотехнологичной продукции, которая обусловлена резким увеличением затрат на создание интеллектуальных систем и переносом центров передовых интеллектуальных разработок из научной сферы в сектор крупного бизнеса, располагающего необходимыми финансовыми и материальными ресурсами для сопровождения данных разработок.

Российский рынок искусственного интеллекта, с учетом общих специфических особенностей создания указанной технологии, обладает тем же выраженным олигополистическим характером, что и глобальный сектор. Однако, после политически мотивированного ухода в 2022 г. крупнейших мировых ИТ-гигантов, данная олигополия приобрела национальные черты: рынок практически поделен между отечественными компаниями: Сбер, Mail.ru и Яндекс. Таким образом, событие неэкономического свойства предоставило российскому ИТ-бизнесу уникальную возможность развиваться, не испытывая на себе негативное влияние конкуренции со стороны глобальных игроков. Вопрос лишь в том, какую стратегию выберут отечественные компании: воспользоваться этой возможностью для реального развития или довольствоваться преференциями олигополистического положения.

Другой негативной тенденцией экономического характера является вывод научно-исследовательских работ из академической среды, что чревато снижением творческой компоненты и подчинением разработок исключительно логике рыночной окупаемости создаваемых технологий. Серьезным

вызовом может стать напряжение на рынке труда на фоне сокращения рабочих мест вследствие автоматизации и роботизации производственных процессов.

В технико-технологической плоскости заметные проблемы возникают в связи с существенным увеличением энергопотребления и сопутствующим ростом углеродного следа, обусловленными экспоненциальным наращиванием вычислительных способностей систем искусственного интеллекта.

Вместе с тем, представляется, что наиболее серьезные риски формируются в этической плоскости, потенциальная уязвимость которой базируется на различных по своей динамике процессах стремительного повышения автономности интеллектуальных систем и практически застойной ситуации в области создания алгоритмов «машинной этики» в указанных системах. Все большую озабоченность вызывает выраженная тенденция к передаче «умным машинам» базовых компетенций в области принятия решений, до настоящего времени являвшихся исключительной прерогативой человека.

Интеллектуальные технологии становятся все сложнее и приобретают все большую автономность от воли человека. Практически бесконтрольный характер их развития, отсутствие установленных границ внедрения искусственного интеллекта чреваты рисками утраты индивидом субъектности по отношению к процессам, протекающим в окружающем мире. Складывающаяся ситуация чревата превращением его самого в объект применения когнитивной деятельности искусственного интеллекта, что предполагает возрастающую актуальность мероприятий по ограничению действия данной технологии этическим нормами. Однако сфера машинной этики, несмотря на значительную проделанную теоретическую работу, до сих пор представляет собой проблемное поле и по степени исследованности пребывает в зачаточном состоянии. Указанное обстоятельство требует серьезного переосмысления существующей парадигмы создания и внедрения в общественное пространство систем искусственного интеллекта в целях минимизации вероятности возникновения этических рисков и угроз для жизнедеятельности человека.

Список источников

1. Гоббс Т. Сочинения. В 2-х т. Т. 1. М.: Мысль, 1989. 622 с. URL: <https://search.rsl.ru/ru/record/01001599658> (дата обращения: 20.11.2023)
2. Декарт Р. Рассуждение о методе, чтобы верно направлять свой разум и отыскивать истину в науках. М.: Академический проект, 2011. 335 с. URL: <https://search.rsl.ru/ru/record/01005034563> (дата обращения: 11.12.2023)

3. Гусев А.В., Добридюк С.Л. Искусственный интеллект в медицине и здравоохранении // Информационное общество. 2017. № 4-5. С. 78–93. EDN: <https://elibrary.ru/yuvvnye>
4. Morgenstern O., Neumann J. Theory of games and economic behavior. New York: Wiley, 1944. 670 p. URL: <https://archive.org/details/theoryofgameseco0000john> (дата обращения: 15.12.2023)
5. Turing A. Computing machinery and intelligence // Mind and language. 1950. Vol. 59. Iss. 236. P. 434–460. URL: https://www.espace-turing.fr/IMG/pdf/Computing_Machinery_and_Intelligence_A-M-_Turing.pdf (дата обращения: 15.01.2024)
6. Moor J.H. Are there decisions computers should never make? // Nature and system. 1979. Vol. 1. Iss. 4. P. 217–229. URL: <https://cse.buffalo.edu/~rapaport/Papers/Papers.by.Others/moor79.pdf> (дата обращения: 15.01.2024)
7. Searle J.R. Minds, brains, and programs // Behavioral and brain sciences. 1980. Vol. 3. Iss. 3. P. 417–424. <https://doi.org/10.1017/S0140525X00005756>
8. Charniak E. Bayesian networks without tears // AI magazine. 1991. Vol. 12. Iss. 4. P. 50–63. URL: http://www.aistudy.com/paper/aaai_journal/AIMag12-04-007.pdf (дата обращения: 10.04.2024)
9. West D., Travis L. The computational metaphor and artificial intelligence: a reflective examination of a theoretical falsehood // AI magazine 1991. Vol. 12. Iss. 1. P. 64–79. URL: http://www.aistudy.com/paper/aaai_journal/AIMag12-01-004.pdf (дата обращения: 12.04.2024)
10. Nadel L. Multiple memory systems: what and why // Journal of cognitive neuroscience. 1992. Vol. 4. Iss. 3. P. 179–188. <https://doi.org/10.1162/jocn.1992.4.3.179>
11. Autor D.H., Levy F., Murnane R.J. The skill content of recent technological change: an empirical exploration // Quarterly journal of economics. 2003. Vol. 118. Iss. 4. P. 1279–1333. <https://doi.org/10.1162/003355303322552801>
12. Zolas N., Kroff Z., Brynjolfsson E., McElheran K., Beede D., Buffington C., Goldschlag N., Foster L., Dinlersoz E. Advanced technologies adoption and use by U.S. firms: evidence from the annual business survey // NBER working paper. Vol. 28290. Cambridge, MA: National Bureau of Economic Research, 2021. 94 p. <https://doi.org/10.3386/w28290>
13. Бурцев М.С., Бухвалов О.Л., Ведяхин А.А. Сильный искусственный интеллект. На подступах к сверхразуму. Москва: Альпина паблишер: Интеллектуальная литература, 2021. 236 с. URL: <https://search.rsl.ru/ru/record/01010578003> (дата обращения: 16.03.2024)
14. Brynjolfsson E., Li D., Raymond L. Generative AI at work // NBER working paper. Vol. 31161. Cambridge, MA: National Bureau of Economic Research, 2023. 65 p. <https://doi.org/10.3386/w31161>
15. Autor D., Chin C., Salomons A.M., Seegmiller B. New frontiers: the origins and content of New Work, 1940–2018 // NBER working paper. Vol. 30389. Cambridge, MA: National Bureau of Economic Research, 2022. 58 p. <https://doi.org/10.3386/w30389>
16. Brynjolfsson E., Rock D., Syverson C. Artificial intelligence and the modern productivity paradox: a clash of expectations and statistics // NBER working paper. Vol. 24001. Cambridge, MA: National Bureau of Economic Research, 2023. 44 p. <https://doi.org/10.3386/w24001>
17. Wallach W., Allen C. Moral machines: teaching robots right from wrong. New York: Oxford University Press, 2009. 294 p. URL: <https://archive.org/details/moralmachinestea0000wall> (дата обращения: 15.01.2024)
18. Weckert J. Intelligent machines, dehumanisation and professional responsibility // In: Computer ethics: philosophical enquiry. Rotterdam: Erasmus University Press, 1997. P. 179–192. <https://doi.org/10.4324/9781315259697>
19. Dreyfus H.L. What computers can't do: the limits of artificial intelligence. New York: Harper Colophon Books, 1979. 374 p. URL: <https://archive.org/details/whatcomputerscan00hube/page/346/mode/2up> (дата обращения: 15.01.2024)
20. Good I.J. Ethical Machines // In: Tenth machine intelligence workshop. Cleveland, Ohio: 1981. P. 5620–5635. URL: <https://courses.cs.washington.edu/courses/cse490e/19wi/readings/good-1980-ethical-machines.pdf> (дата обращения: 16.01.2024)
21. Moor J.H. The nature, importance, and difficulty of machine ethics // IEEE intelligent systems. 2006. Vol. 21. Iss. 4. P. 18–21. <https://doi.org/10.1109/MIS.2006.80>
22. Бостром Н. Искусственный интеллект: этапы, угрозы, стратегии; пер. с англ. Москва: Манн, Иванов и Фербер, 2016. 490 с. URL: <https://search.rsl.ru/ru/record/01008120827> (дата обращения: 16.01.2024)

23. Russell S. Human compatible: Artificial intelligence and the problem of control. New York: Viking, 2019. 308 p. URL: <https://searchworks.stanford.edu/view/13513564> (дата обращения: 10.01.2024)
24. Luccioni A., Viguier S., Ligozat A.-L. Estimating the carbon footprint of BLOOM a 176B parameter language model // *Journal of machine learning research*. 2023. Iss. 11. P. 1–15. <https://doi.org/10.48550/arXiv.2211.02001>
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, July 28-August 2, 2019. P. 3645–3650. <https://doi.org/10.48550/arXiv.1906.02243>
26. Acemoglu D., Restrepo P. Automation and new tasks: how technology displaces and reinstates labor // *Journal of economic perspectives*. 2019. Vol. 33. Iss. 2. P. 3–30. <https://doi.org/10.1257/jep.33.2.3>
27. Malle B.F., Scheutz M., Arnold Th., Voiklis J., Cusimano C. Sacrifice one for the good of many?: People apply different moral norms to human and robot agents // In: HRI '15: Proceedings of the Tenth annual ACM/IEEE international conference on human-robot interaction. March 2015. P. 117–124. <https://doi.org/10.1145/2696454.2696458>

Статья поступила в редакцию 15.04.2024; одобрена после рецензирования 27.06.2024; принята к публикации 28.06.2024

Об авторе:

Щербаков Геннадий Анатольевич, доктор экономических наук, профессор, профессор кафедры «Системный анализ в экономике»; SPIN-код: 7096-0093, Scopus ID: 57212108593

Автор прочитал и одобрил окончательный вариант рукописи.

References

1. Hobbes T. Works in 2 vol. Vol. 1. Moscow: Mysl, 1989. 622 p. URL: <https://search.rsl.ru/ru/record/01001599658> (accessed: 20.11.2023) (In Russ.)
2. Descartes R. Discourse on the Method to Correctly Direct your Mind and Find the Truth in the Sciences. Moscow: Academic Project, 2011. 335 p. URL: <https://search.rsl.ru/ru/record/01005034563> (accessed: 11.12.2023) (In Russ.)
3. Gusev A.V., Dobridnyuk S.L. Artificial intelligence in medicine and healthcare. *Information Society*. 2017; (4-5):78–93. EDN: <https://elibrary.ru/yuvyne> (In Russ.)
4. Morgenstern O., Neumann J. Theory of games and economic behavior. New York: Wiley, 1944. 670 p. URL: <https://archive.org/details/theoryofgameseco0000john> (accessed: 15.12.2023) (In Eng.)
5. Turing A. Computing machinery and intelligence. *Mind and language*. 1950; (59(236)):434–460. URL: https://www.espace-turing.fr/IMG/pdf/Computing_Machinery_and_Intelligence_A-M-Turing.pdf (accessed: 15.01.2024) (In Eng.)
6. Moor J.H. Are there decisions computers should never make? *Nature and system*. 1979; (1(4)):217–229. URL: <https://cse.buffalo.edu/~rapaport/Papers/Papers.by.Others/moor79.pdf> (accessed: 15.01.2024) (In Eng.)
7. Searle J.R. Minds, brains, and programs. *Behavioral and brain sciences*. 1980; (3(3)):417–424. <https://doi.org/10.1017/S0140525X00005756> (In Eng.)
8. Charniak E. Bayesian networks without tears. *AI magazine*. 1991; (12(4)):50–63. URL: http://www.aistudy.com/paper/aaai_journal/AIMag12-04-007.pdf (accessed: 10.04.2024) (In Eng.)
9. West D., Travis L. The computational metaphor and artificial intelligence: a reflective examination of a theoretical falsework. *AI magazine*. 1991; (12(1)):64–79. URL: http://www.aistudy.com/paper/aaai_journal/AIMag12-01-004.pdf (accessed: 12.04.2024) (In Eng.)
10. Nadel L. Multiple memory systems: what and why. *Journal of cognitive neuroscience*. 1992; 4(3):179–188. <https://doi.org/10.1162/jocn.1992.4.3.179> (In Eng.)
11. Autor D., Levy F., Murnane R. The skill content of recent technological change: an empirical exploration. *Quarterly journal of economics*. 2003; (118(4)):1279–1333. <https://doi.org/10.1162/003355303322552801> (In Eng.)
12. Zolas N., Kroff Z., Brynjolfsson E., McElheran K., Beede D., Buffington C., Goldschlag N., Foster L., Dinlersoz E. Advanced technologies adoption and use by U.S. firms: evidence from the annual business survey. *NBER working*

paper. Iss. 28290. Cambridge, MA: National Bureau of Economic Research, 2021. 94 p. <https://doi.org/10.3386/w28290> (In Eng.)

13. Burtsev M.S., Bukhvalov O.L., Vedyakhin A.A. Strong artificial intelligence. On the approaches to supermind. Moscow: Alpina Publisher: Intellectual Literature, 2021. 236 p. URL: <https://search.rsl.ru/ru/record/01010578003> (accessed: 16.03.2024) (In Russ.)

14. Brynjolfsson E., Li D., Raymond L. Generative AI at work. *NBER working paper*. Iss. 31161. Cambridge, MA: National Bureau of Economic Research, 2023. 65 p. <https://doi.org/10.3386/w31161> (In Eng.)

15. Autor D., Chin C., Salomons A.M., Seegmiller B. New frontiers: the origins and content of New Work, 1940–2018. *NBER working paper*. Iss. 30389. Cambridge, MA: National Bureau of Economic Research, 2022. 58 p. <https://doi.org/10.3386/w30389> (In Eng.)

16. Brynjolfsson E., Rock D., Syverson C. Artificial intelligence and the modern productivity paradox: a clash of expectations and statistics. *NBER working paper*. Iss. 24001. Cambridge, MA: National Bureau of Economic Research, 2023. 44 p. <https://doi.org/10.3386/w24001> (In Eng.)

17. Wallach W., Allen C. Moral machines: teaching robots right from wrong. New York: Oxford University Press, 2009. 294 p. URL: <https://archive.org/details/moralmachinestea0000wall> (accessed: 15.01.2024) (In Eng.)

18. Weckert J. Intelligent machines, dehumanisation and professional responsibility. In: *Computer ethics: philosophical enquiry*. Rotterdam: Erasmus University Press, 1997. P. 179–192. <https://doi.org/10.4324/9781315259697> (In Eng.)

19. Dreyfus H.L. What computers can't do: the limits of artificial intelligence. New York: Harper Colophon Books, 1979. 374 p. URL: <https://archive.org/details/whatcomputerscan00hube/page/346/mode/2up> (accessed: 15.01.2024) (In Eng.)

20. Good I.J. Ethical machines. In: *Tenth machine intelligence workshop*. Cleveland, Ohio: 1981. P. 5620–5635. URL: <https://courses.cs.washington.edu/courses/cse490e/19wi/readings/good-1980-ethical-machines.pdf> (accessed: 16.01.2024) (In Eng.)

21. Moor J.H. The nature, importance, and difficulty of machine ethics. *IEEE intelligent systems*. 2006; (21(4)):18–21. <https://doi.org/10.1109/MIS.2006.80> (In Eng.)

22. Bostrom N. Superintelligence: Paths, dangers, strategies. New York: Oxford University Press, 2014. 328 p. URL: <https://dorshon.com/wp-content/uploads/2017/05/superintelligence-paths-dangers-strategies-by-nick-bostrom.pdf> (accessed: 16.01.2024) (In Eng.) (Russ. ed.: Bostrom N. Artificial intelligence: stages, threats, strategies. Moscow: Mann, Ivanov and Ferber, 2016. 490 p.)

23. Russell S. Human compatible: Artificial intelligence and the problem of control. New York: Viking, 2019. 308 p. URL: <https://searchworks.stanford.edu/view/13513564> (accessed: 10.01.2024) (In Eng.)

24. Luccioni A., Viguier S., Ligozat A.-L. Estimating the carbon footprint of BLOOM, a 176B parameter language model. *Journal of machine learning research*. 2023; (11):1–15. <https://doi.org/10.48550/arXiv.2211.02001> (In Eng.)

25. Strubell E., Ganesh A., McCallum A. Energy and policy considerations for deep learning in NLP. In: *Proceedings of the 57th annual meeting of the Association for computational linguistics*. Florence, Italy, July 28–August 2, 2019. P. 3645–3650. <https://doi.org/10.48550/arXiv.1906.02243> (In Eng.)

26. Acemoglu D., Restrepo P. Automation and new tasks: How technology displaces and reinstates labor. *Journal of economic perspectives*. 2019; (33(2)):3–30. <https://doi.org/10.1257/jep.33.2.3> (In Eng.)

27. Malle B.F., Scheutz M., Arnold Th., Voiklis J., Cusimano C. Sacrifice one for the good of many?: People apply different moral norms to human and robot agents. In: *HRI '15: Proceedings of the Tenth annual ACM/IEEE international conference on human-robot interaction*. March 2015. P. 117–124. <https://doi.org/10.1145/2696454.2696458> (In Eng.)

The article was submitted 15.04.2024; approved after reviewing 27.06.2024; accepted for publication 28.06.2024

About the author:

Gennady A. Shcherbakov, Doctor of Economics Sciences, Professor; SPIN: 7096-0093, Scopus ID: 57212108593

The author read and approved the final version of the manuscript.