

Оригинальная статья / Original article

УДК 338.24

<https://doi.org/10.21869/2223-1552-2024-14-5-60-73>



Влияние искусственного интеллекта на экономическое развитие: этические аспекты внедрения и разработки технологий

И. Ф. Мальцева^{1✉}, Е. С. Сазонова¹

¹ Юго-Западный государственный университет
ул. 50 лет Октября, д. 94, г. Курск 305040, Российская Федерация

✉ e-mail: irina-fedorovna@inbox.ru

Резюме

Актуальность. Цифровизация является одним из главных приоритетов экономики. Технологии машинного обучения и искусственного интеллекта внедряются передовыми компаниями в производственные и управленческие процессы, меняя характер деятельности и способы ведения бизнеса. Возникают новые вопросы, связанные с законодательным регулированием технологий, распределением зон ответственности между человеком и роботизированными системами, появляются правовые отношения, не предусмотренные нормативными рамками. Технологическое развитие ведет также к глубокой трансформации общества. Новые механизмы взаимодействия порождают ряд этических вопросов. В связи с этим тема исследований внедрения искусственного интеллекта становится все более популярной в научной среде.

Цель статьи – изучить особенности внедрения цифровых систем, в том числе на основе искусственного интеллекта, а также систематизировать этические проблемы использования таких технологий.

Задачи: проанализировать нормативно-правовую базу регулирования систем искусственного интеллекта; определить базовые принципы его разработки и внедрения; рассмотреть основные проблемы, возникающие в связи с расширением сферы применения искусственного интеллекта, а также варианты их решения.

Методология. Для достижения поставленной цели в исследовании применялись методы системного анализа и синтеза, библиографического анализа. Информационной базой исследования послужили труды отечественных и зарубежных ученых, посвященные ИИ, нормативная база федерального и регионального уровня, материалы научных журналов и конференций.

Результаты. В статье приведен литературный обзор современных научных исследований, посвященных этическим вопросам внедрения ИИ, рассмотрены основные нормативные акты, применяемые как в международной практике, так и разработанные в России.

Выводы. На основе проведенного анализа представлена систематизация принципов и этических проблем внедрения систем искусственного интеллекта. При имеющихся преимуществах использования новой технологии необходимо создавать законодательную и институциональную основу, способствующую минимизации рисков и содействующую достижению целей устойчивого развития.

Ключевые слова: искусственный интеллект; машинное обучение; этические проблемы; цифровая трансформация; законодательное регулирование.

Финансирование: Работа выполнена в рамках реализации программы развития ФГБОУ ВО «Юго-Западный государственный университет» проекта «Приоритет – 2030».

Конфликт интересов: В представленной публикации отсутствует заимствованный материал без ссылок на автора и (или) источник заимствования, нет результатов научных работ, выполненных авторами публикации лично и (или) в соавторстве, без соответствующих ссылок. Авторы декларируют отсутствие конфликта интересов, связанных с публикацией данной статьи.

© Мальцева И. Ф., Сазонова Е. С., 2024

Для цитирования: Мальцева И. Ф., Сазонова Е. С. Влияние искусственного интеллекта на экономическое развитие: этические аспекты внедрения и разработки технологий // Известия Юго-Западного государственного университета. Серия: Экономика. Социология. Менеджмент. 2024. Т. 14, № 5. С. 60–73. <https://doi.org/10.21869/2223-1552-2024-14-5-60-73>

Поступила в редакцию 17.08.2024

Принята к публикации 11.09.2024

Опубликована 31.10.2024

The impact of artificial intelligence on economic development: ethical aspects of technology implementation and development

Irina F. Maltseva¹✉, Elizaveta S. Sazonova¹

¹ Southwest State University
50 Let Oktyabrya Str. 94, Kursk 305040, Russian Federation

✉ e-mail: irina-fedorovna@inbox.ru

Abstract

Relevance. Digitalization is one of the main priorities of the economy. Machine learning and artificial intelligence technologies are introduced by advanced companies into production and management processes, and change the nature of their activities and the way they do business. New issues related to the legislative regulation of technologies, the distribution of areas of responsibility between humans and robotic systems arise, and legal relations that are not provided for by the regulatory framework appear. Technological development also leads to a profound transformation of society. New mechanisms of interaction raise a number of ethical issues. In this regard, the topic of research on the introduction of artificial intelligence is becoming increasingly popular in the scientific community.

The purpose of the article is to study the features of the introduction of digital systems, including those which are based on artificial intelligence, as well as to systematize the ethical problems of using such technologies.

Objectives: analyze the regulatory framework for the regulation of artificial intelligence systems, determine the basic principles of its development and implementation, consider the main problems arising in connection with the expansion of the scope of artificial intelligence, as well as options for their solutions.

Methodology. To achieve the goal in the research are used methods of system analysis and synthesis, bibliographic analysis. The works of domestic and foreign scientists devoted to AI, the regulatory framework of the federal and regional levels, materials of scientific journals and conferences have become the information base of the study.

Results. A literary review of modern scientific research on ethical issues of implementation is presented in the article, the main normative acts used both in international practice and developed in Russia are considered.

Conclusions. The systematization of the principles and ethical problems of the introduction of artificial intelligence systems based on the analysis are presented. The use of new technology has advantages, but it is necessary to create a legislative and institutional framework that will minimize risks and achieve sustainable development goals.

Keywords: artificial intelligence; machine learning; ethical issues; digital transformation; legislative regulation.

Funding: The work was carried out within the framework of the implementation of the development program of the Southwest State University of the Priority 2030 project.

Conflict of interest: In the presented publication there is no borrowed material without references to the author and (or) source of borrowing, there are no results of scientific works performed by the authors of the publication, personally and (or) in co-authorship, without relevant links. The authors declares no conflict of interest related to the publication of this article.

For citation: Maltseva I.F., Sazonova E.S. The impact of artificial intelligence on economic development: ethical aspects of technology implementation and development. *Izvestiya Yugo-Zapadnogo gosudarstvennogo universiteta. Seriya: Ekonomika. Sotsiologiya. Menedzhment = Proceedings of the Southwest State University. Series: Economics, Sociology and Management.* 2024;14(5):60–73. (In Russ.) <https://doi.org/10.21869/2223-1552-2024-14-5-60-73>

Received 17.08.2024

Accepted 11.09.2024

Published 31.10.2024

Введение

Ранние исследования в области искусственного интеллекта начались практически сразу после создания первых вычислительных машин. В 1954 г. американский учёный Аллен Ньюэлл решил разработать программу для игры в шахматы, а голландские психологи, включая А. Де Гроота, присоединились к проекту и изучали стиль игры известных шахматистов. Через два года совместной работы команда создала язык программирования IPL (Information Processing Language), ставший первым символьным языком обработки списков. Затем была создана первая программа, которая считается важным достижением в области искусственного интеллекта, – «Логик-теоретик» (1956), она предназначалась для автоматического доказательства теорем в исчислении высказываний [1; 2].

В середине 1970-х годов начали появляться первые прикладные интеллектуальные системы, которые использовали разные методы представления знаний для решения задач. Эти системы стали известны как экспертные системы.

Развитие искусственного интеллекта (ИИ) подталкивает исследователей к необходимости ответа на ряд философских вопросов: какова природа интеллекта и что именно может быть воспроизведено в компьютерных системах (какие процессы нашего мозга могут быть воспроизведены, а какие процессы остаются недоступными, какие типы знаний и умений являются необходимыми для эмоций, мотивации и желаний, с какими этическими проблемами предстоит столкнуться в связи с внедрением ИИ и какие меры по их смягчению должны быть приоритетными для их решения) [3]. Отдельно стоит вопрос о том, возможно ли сознание в компьютерной системе [4]. Дуализм, предложенный Рене Декартом в XVII в., предполагает, что разум и материя представляют собой отдельные сущности, взаимодействующие друг с дру-

гом. Однако многие философы оспаривают эту точку зрения, утверждая, что разум и тело являются взаимозависимыми и неделимыми составляющими [5].

Определение ИИ, обозначенное в рамках Дартмутского летнего исследовательского проекта по искусственному интеллекту 1955 г. [6], предполагало, что каждый аспект человеческого интеллекта при достаточно подробном описании может быть смоделирован машиной: рассуждение, обучение, решение проблем, принятие решений и креативность [7]. При этом основной акцент был сделан на вычислительной мощности и возможностях ИИ, поскольку знания о природе мышления, интеллекта и работе мозга пока ограничены.

Ричард Фейнман, нобелевский лауреат и физик, пришёл к выводу, что великие открытия не могут быть сделаны путём логических рассуждений. ИИ не способен совершать абсолютно новые открытия, поскольку творческий процесс требует интуитивного понимания и нестандартного мышления. Только человеческий мозг, функционирующий согласно природным законам, способен на творческие озарения и создание уникальной картины мира [8].

Другое мнение на этот счет приведено в исследовании А. Barcaui, A. Monat [9], где употребляется такое понятие, как генеративный ИИ, – относится к отрасли ИИ, которая использует алгоритмы и методы машинного обучения для генерации новых, креативных решений или оптимизации существующих процессов.

Одним из наиболее значимых достижений в области генеративного ИИ является разработка больших языковых моделей (LLM). Эти модели воплощают способность ИИ понимать, генерировать и трансформировать человеческий язык. В категории генеративного искусственного интеллекта значительную роль играет генеративный предварительный предобученный трансформер (GPT). Это языковая модель, разработанная OpenAI, кото-

рая может быть применена для создания различного контента: текст, аудио, изображения и видео [10].

Хотя генеративный ИИ предоставляет уникальные возможности для управления проектами, принятие решений человеком остается неотъемлемой частью успеха проектов. A. Nieto-Rodriguez, R.V. Vargas приводят статистику, согласно которой сегодня только 35% проектов завершаются успешно, при этом одной из причин такого показателя является низкий уровень зрелости технологий, доступных для управления проектами. Авторы предполагают, что ситуация находится на пороге перемен, поскольку исследователи, стартапы и инновационные организации начинают применять ИИ, машинное обучение и другие передовые технологии для управления проектами [11].

В использовании генеративного ИИ можно выделить 3 уровня.

1. Стратегический уровень:

– принятие решений: менеджеры могут применять чат GPT при принятии стратегических решений, используя рекомендации в конкретных случаях;

– менеджмент знаний: организации могут применять чат GPT для хранения, преобразования и распространения организационных данных.

2. Функциональный уровень:

– обслуживание клиентов: чат GPT можно использовать для автоматизации взаимодействия с клиентами, например, для ответов на часто задаваемые вопросы и предоставления информации;

– управление персоналом: чат GPT можно использовать для автоматизации и оптимизации различных процессов управления персоналом, например, просмотр резюме, предоставление обратной связи кандидатам, прием на работу новых сотрудников или реагирование на просьбы сотрудников.

3. Административный уровень. Автоматизация типовых повторяющихся задач: чат GPT можно использовать для планирования встреч, бронирования по-

ездок, управления календарями, создания шаблонов деловых писем, составления отчетности [10].

Менеджеры проектов обладают рядом навыков и качеств, которые трудно воспроизвести с помощью одних только систем ИИ. Принятие нестандартных решений, креативность и навыки межличностного общения входят в число ценных качеств, которые может принести в работу только человеческий интеллект. Эти способности особенно важны в сложных и динамичных проектных средах, где адаптивность, интуиция и социальный интеллект играют важную роль. Тем не менее значительные экономические и социальные выгоды, которые создает ИИ, сопровождаются опасениями, что эти технологии могут оказать пагубное воздействие на права человека [12]. Проблемы охватывают различные аспекты: от защиты данных и информационной безопасности до алгоритмических искажений, безработицы, усиления экономического неравенства и манипулирования демократическими процессами [13].

Стремительное развитие ИИ приводит к возникновению новых проблем этики и прав человека, которые часто становятся темой общественных и научных дебатов [14; 15].

Материалы и методы

Алгоритмы ИИ могут справиться с несколькими возникающими проблемами за гораздо меньшее время, чем человек, но в то же самое время алгоритмы создают и новые проблемы. Как элементы, принадлежащие к миру кода, они подвержены ошибкам. По скорости вычислений могут значительно ускорить человеческие ошибки и мгновенно внедрять в жизнь предрассудки и стереотипы своих программистов и тренеров нейронных сетей.

В этом процессе важно проанализировать эти новые типы проблем и их взаимодействие в алгоритмически ускоренном мире. Только те, кто понимает работу ал-

горитмов, могут увеличить их преимущества и уменьшить их вредный потенциал.

Теоретическую основу для выявления и систематизации рисков и этических проблем, связанных с разработкой и внедрением систем ИИ, составили труды российских и зарубежных ученых, доклады международных организаций и общественных объединений, в частности ОЭСР и ЮНЕСКО.

Для определения и выявления общих этических принципов был использован метод сравнительного анализа, систематизированы нормативно-правовые документы, регулирующие развитие ИИ как на международном уровне, так и в отдельных странах.

Результаты и их обсуждение

В связи с тем, что ИИ уже основательно вошел в обиход нашей повседневной жизни, стали возникать нетехнические проблемы, связанные с использованием его механизмов. Их часто описывают как этические проблемы [16].

Рассмотрим основные риски и проблемы, связанные с разработкой и внедрением систем ИИ:

1. Проблема отсутствия нейтральности. Она представляется особенно важной, так как алгоритмы обычно не являются нейтральными инструментами. Их создает человек, и таким образом они часто отражают убеждения своего создателя. В случае с ИИ мы имеем дело с отражением предрассудков, содержащихся в данных, на которых обучалась нейронная сеть.

Эта проблема проявляется по-разному. В коде многих ИИ проявляются социально-негативные явления. Например, в 2015 г. Amazon обнаружил, что созданный корпорацией алгоритм покровительствовал мужчинам при поиске подходящих кандидатов. Не потому, что мужчины были объективно более подходящими кандидатами, а только потому, что алгоритм был обучен предыдущими данными, собранными сотрудниками отдела кадров, которые предпочитают мужчин.

Amazon обнаружил проблему, потому что им нужны были хорошие сотрудники, а выбор алгоритма сужал список потенциальных кандидатов. В результате система была пересмотрена, хотя сомнений по-прежнему осталось достаточно для того, чтобы программа была закрыта в конце 2017 г.

Rigotti C., Fosch-Villaronga E. в процессе анализа дискриминации при подборе персонала с помощью ИИ приходят к выводу о необходимости гарантии справедливости приложений ИИ в процессе найма [17].

Исследование ЮНЕСКО, посвященное анализу языковых моделей, также показало, что ИИ подвержен дискриминации и искажению информации, содержит гендерные и иные стереотипы, неверно отражает культурные и иные традиции стран. Это может быть следствием как вводных данных, на которых обучается модель, так и непропорциональным представлением тех или иных групп в исследовании [18].

2. Риски, связанные с трансформацией рынка труда. Внедрение ИИ оказывает влияние на способы организации бизнес-процессов и может усилить существующую дифференциацию в обществе, в т. ч. за счет неравного доступа к технологиям ИИ. Акцент на этическом аспекте трансформации рабочих мест столь же важен, как и на технических аспектах [19; 20].

3. Риски безопасности, конфиденциальности и защиты данных [21]. ИИ может генерировать контент на опасные темы, создавая угрозу безопасности жизни и здоровью людей. При правильном запросе модели ИИ могут раскрыть персональные или платежные данные, на которых обучалась система.

Этические проблемы приобретают особую значимость, когда касаются жизни и здоровья людей. Неслучайно поэтому многие исследователи, занимающиеся вопросами медицины, уделяют внимание данному вопросу. Так, в своем исследовании Chauhan C., Gullapalli R. при изу-

чении влияния ИИ и систем принятия решений помимо перечисленных рисков отмечают также вопрос о том, надо ли информировать, что решение о диагнозе принималось с помощью ИИ, стоит ли брать согласие на применение данной технологии у пациента и др. Авторы приходят к выводам о том, что ИИ в медицине должен отвечать трем принципам: прозрачность, подотчетность и управление [22].

4. Проблема деградации социальных взаимодействий и риск воздействия в долгосрочной перспективе на когнитивные способности, связанные с грамотностью, включая письмо, понимание и критическое мышление.

Большой блок вопросов здесь связан с использованием ИИ в науке и образовании [23]. Понимая невозможность тотального запрета на применение студентами ИИ при написании работ, некоторые вузы (МГПУ, НИУ ВШЭ, САФУ, ТюмГУ) официально разрешили его использование в качестве инструмента систематизации информации и поиска новых идей, генерации изображений и помощи в структурировании и оформлении работ.

Дискуссия также сосредоточилась на возможности использования ChatGPT для мошенничества при оценках; об авторстве и ссылках на научные исследования; о деградации социальных отношений в образовательном процессе [18].

5. Риски надежности генерируемой информации. ИИ может обманывать и выдавать вымышленную информацию, которая понравится пользователю. Некоторые системы искусственного интеллекта научились обманывать тесты, предназначенные для оценки их безопасности. В работе Park P., Goldstein S. и др. приводятся примеры подобного поведения систем ИИ [24].

6. Проблема распределения ответственности. При работе с ИИ важно понимать, как будет распределяться ответственность за принятые с участием ИИ решения. В России в соответствии с кодексом этики в сфере ИИ ответствен-

ность за принятие решений всегда лежит на человеке [25].

7. Риск использования ИИ для манипулирования людьми. Сегодня нейроданные, собранные с помощью устройств, становятся отдельным продуктом. ЮНЕСКО отмечает опасность при использовании устройств, которые собирают нейроданные и могут быть применены для манипулирования людьми, они представляют угрозу для психической неприкосновенности и свободы мысли [26].

Приведенные риски можно объединить в 4 взаимосвязанных блока:

1. Этика прав человека.
2. Моральная этика.
3. Информационная этика.
4. Этика ответственности.

Для минимизации обозначенных рисков и проблем, регулирования процессов разработки и внедрения ИИ формируются соответствующая нормативно-правовая база, рекомендации и кодексы. В последние годы многие международные организации стали принимать документы, относящиеся к «мягкому праву», которые носят рекомендательный характер. Эти документы направлены на определение главных принципов и норм, регулирующих разработку и использование систем ИИ как государственными, так и частными структурами.

Одним из первых стандартов стали рекомендации ОЭСР по ИИ [27]. В 2020 г. Европейская комиссия опубликовала «Белую книгу по ИИ – европейский подход к совершенству и доверию» (A White paper on Artificial Intelligence – A European approach to excellence and trust) [28]. В ней отражены все действия, позволяющие достичь двойной цели – стимулировать внедрение ИИ в жизнь людей и одновременно с этим как можно больше устранить риски использования этих технологий. В 2024 г. был разработан регламент ЕС об искусственном интеллекте, основным нововведением которого стала дифференциация систем ИИ и требований к ним по уровню риска [29].

На Генеральной конференции ЮНЕСКО в 2021 г. 193 странами было принято первое глобальное соглашение по этике ИИ – Рекомендации по этическим аспектам ИИ, которые должны стать этическими ориентирами и нормативной основой для строгого соблюдения верховенства права в цифровой среде. Основная цель – способствовать развитию ИИ, но при этом не допускать нанесения ущерба отдельным людям и экосистемам.

В отдельных странах также принимаются документы, регулирующие ИИ. Наиболее проработаны механизмы государственного нормативно-правового обеспечения разработки и развития ИИ в Китае. Filirova I. приводит детальный обзор и эволюцию документов регулирования вопросов ИИ в Китае, в т. ч. связанных с этической стороной [30].

В октябре 2021 г. РФ был разработан Кодекс этики в сфере искусственного интеллекта [25], который включает в себя единую систему рекомендательных принципов и правил, которые предназначены для создания среды доверенного развития технологий искусственного интеллекта. Кодекс был подписан многими компаниями, включая *Сбербанк, Яндекс, МТС, VK, Газпром нефть* и *Российский фонд прямых*

инвестиций, а также ведущие научно-исследовательские организации. Всего на сегодняшний день документ подписали более 250 компаний.

Отдельные документы по регулированию ИИ приняты также в ОАЭ, Сингапуре, Японии, Южной Корее и др. Продолжается работа на международном уровне: в настоящее время ЮНЕСКО ведет разработку первого глобального стандарта по этическим аспектам нейротехнологий, утверждение которого ожидается к концу 2025 г.

В таблице 1 представлены документы, регулирующие разработку и использование систем искусственного интеллекта в сфере этики как на международном уровне, так и на уровне отдельных стран.

Из таблицы 1 видно, что во многом принципы совпадают, однако стоит отметить, что зачастую они носят рекомендательный характер, при этом достаточно сложно организовать контроль и мониторинг соблюдения данных принципов. Первое исследование прогресса в направлении ответственного отношения к ИИ показало, что он значительно отстает от динамики разработки и внедрения ИИ, что создает достаточную основу для беспокойства [35].

Таблица 1. Законодательное регулирование развития ИИ

Объединение / страна	Название законодательного акта	Этические принципы разработки и внедрения ИИ
ОЭСР	Принципы ОЭСР по искусственному интеллекту [27]	Инклюзивный рост, устойчивое развитие и благополучие Ценности и справедливость, ориентированные на человека Прозрачность и объяснимость Надежность, защита и безопасность Ответственность
ЮНЕСКО	Рекомендации по этике искусственного интеллекта [31]	Отсутствие вреда для человека Защита и безопасность Справедливость и недискриминация Устойчивое развитие Право на конфиденциальность и защиту данных Управляемость и мониторинг, приоритет принятия важных решений человеком Прозрачность и объяснимость Ответственность и подотчетность Информированность и грамотность Многостороннее и адаптивное управление и сотрудничество

Объединение / страна	Название законодательного акта	Этические принципы разработки и внедрения ИИ
США	Указ о безопасном, надежными заслуживающем доверия ИИ [32]	Защита конфиденциальности Продвижение справедливости и гражданских прав Защита потребителей, пациентов и студентов Поддержка работников Содействие инновациям и конкуренции Продвижение американского лидерства за границей Обеспечение ответственного и эффективного использования ИИ правительством
Канада	Закон об искусственном интеллекте и данных [33]	Управляемость и мониторинг Прозрачность Справедливость и равноправие Безопасность Отчетность Достоверность и надежность
Великобритания	Внедрение принципов регулирования ИИ в Великобритании [34]	Безопасность, защищенность от взлома, надежность Надлежащая прозрачность и объяснимость Справедливость Подотчетность и управление Возможность оспаривания выходных данных от ИИ и возмещения ущерба
Китай	План развития искусственного интеллекта нового поколения Пекинские принципы искусственного интеллекта Кодекс этики искусственного интеллекта нового поколения [30]	Рост благосостояния человека Содействие справедливости и правосудию Защита конфиденциальности и безопасности Управляемость и надежность Ответственность Повышение этической грамотности
Россия	Национальная стратегия «О развитии искусственного интеллекта в Российской Федерации» ¹ Кодекс этики в сфере искусственного интеллекта [25]	Защита прав и свобод человека Безопасность Прозрачность Технологический суверенитет Целостность инновационного цикла Наиболее эффективное использование технологий искусственного интеллекта Поддержка конкуренции Открытость и доступность Преемственность Защищенность Достоверность исходных данных

¹ О развитии искусственного интеллекта в Российской Федерации: Указ Президента Российской Федерации от 10.10.2019 г. № 490. URL: <http://www.kremlin.ru/acts/bank/44731/page/1> (дата обращения: 26.07.2024).

Если разобрать и систематизировать представленные принципы ИИ в разных странах мира, то можно выделить следующие:

- ответственность;
- конфиденциальность;

- безопасность;
- человекоориентированный и гуманистический подход.

На рисунке 1 представлены основные этические проблемы, принципы ИИ, уровни регулирования и цели.



Рис. 1. Этические проблемы, принципы ИИ, уровни их регулирования и цели

Выводы

Технологии ИИ имеют большое влияние на нашу жизнь, но, улучшая её, они одновременно могут и нарушать права и свободу человека. Приведенный обзор рисков, изучение научных исследований в сфере этики разработки и внедрения систем ИИ позволили прийти к следующим выводам:

1. Для успешного развития ИИ необходимы согласованные действия трех сторон «государство – общество – бизнес»:

1) государство. Для обеспечения целей устойчивого развития и минимизации

обозначенных рисков необходима детальная проработка законодательной базы, совершенствование систем мониторинга и контроля, проведение политики, обеспечивающей технологический суверенитет и лидерство в области ИИ;

2) общество. Важным является повышение доступности информации об использовании ИИ в разных сферах (образование, медицина и т. д.) и повышение грамотности населения с целью формирования ответственного отношения к использованию новой технологии, снижения уровня сопротивления при внедрении систем ИИ;

3) бизнес. Содействие распространению ИИ в экономике (сфера услуг, сельское хозяйство, промышленность, управление) с целью роста производительности труда в условиях кадрового дефицита. Одним из способов решения может стать создание набора типовых решений ИИ для бизнеса (например, типовые карты пути сотрудников, показывающие, на каком этапе могут быть применены те или иные технологии ИИ), вторым – разработка соответствующих федеральных платформ для бизнеса на базе ИИ (частично такая работа уже ведется).

2. Внедрение ИИ призвано улучшить благосостояние общества и содействовать достижению целей устойчивого раз-

вития, однако следует уделять большое внимание соблюдению прав человека и этических принципов. Разделяя системы ИИ по уровню риска, государство может разрабатывать механизмы, которые, с одной стороны, будут стимулировать развитие технологий ИИ, с другой – обеспечивать необходимый контроль за высокорисковыми системами ИИ, представляющими угрозу жизни и здоровью людей, национальной безопасности.

3. Технологии ИИ не ограничиваются границами стран, поэтому необходимо укрепление международного сотрудничества в фундаментальных вопросах разработки и распространения технологий ИИ.

Список литературы

1. Новиков Ф. А. Символический искусственный интеллект: математические основы представления знаний. М.: Юрайт, 2024. 278 с.
2. Поспелов Д. А. Моделирование рассуждений. Опыт анализа мыслительных актов. М.: Радио и связь, 1989. 184 с.
3. Exploring ethics and human rights in artificial intelligence – A Delphi study / В. С. Stahl [et al.] // *Technol Forecast Soc Change*. 2023. Vol. 191. P. 122502. <https://doi.org/10.1016/j.techfore.2023.122502>
4. Романов Р. В. Классические концепции решения проблем искусственного интеллекта (философский аспект) // *Преподаватель XXI век*. 2012. № 3. С. 241–247.
5. Канышева О. А. Учение Рене Декарта о душе и сознании // *Вестник Воронежского государственного университета. Серия: Философия*. 2021. № 3. С. 99–107.
6. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence / J. McCarthy [et al.] // *AI Magazine*. 2006. N. 27(4).
7. Rai A., Sarker S. Editor's Comments Next-Generation Digital Platforms: Toward Human-AI Hybrids // *MIS Quarterly*. 2019. Vol. 43, N 1. P. 1–7.
8. Синельникова А. П. Сможет ли искусственный интеллект заменить человека? // *Вестник науки*. 2020. № 2(23). С. 77–79.
9. Barcaui A., Monat A. Who is better in project planning? Generative artificial intelligence or project managers? // *Project Leadership and Society*. 2023. Vol. 4. P. 100101. <https://doi.org/10.1016/j.plas.2023.100101>
10. Generative artificial intelligence as a new context for management theories: analysis of ChatGPT / P. Korzynski [et al.] // *Central European Management Journal*. 2023. Vol. 31, N 1. P. 0091. <https://doi.org/10.1108/CEMJ-02-2023-0091>
11. Nieto-Rodriguez A., Vargas R. How AI Will Transform Project Management // *PM World Journal*. 2023. P. 1–2. URL: <https://pmworldlibrary.net/wp-content/uploads/2023/02/230224-hbr-how-ai-will-transform-project-management-publishing-news.pdf> (дата обращения: 27.07.2024).

12. Mantelero A., Esposito M. S. An evidence-based methodology for human rights impact assessment (HRIA) in the development of AI data-intensive systems // *Computer Law & Security Review*. 2021. Vol. 41. P. 105561. <https://doi.org/10.1016/j.clsr.2021.105561>
13. 2024 Регламент Европейского Союза об искусственном интеллекте, АНО «Цифровая экономика». URL: https://ai.gov.ru/knowledgebase/dokumenty-po-razvitiyu-ii-v-drugikh-stranakh/2024_reglament_evropeyskogo_soyuza_ob_iskusstvennom_intellekte_ano_cifrovaya_e_konomika/ (дата обращения: 24.07.2024).
14. Lance B. Eliot. Responsible AI Relishes Preeminent Boost Via AI Ethics Proclamation By Top Professional Society The ACM // *Forbes*. URL: <https://www.forbes.com/sites/lanceeliot/2022/11/27/responsible-ai-relishes-mighty-boost-via-ai-ethics-proclamation-rolled-out-by-esteemed-computing-profession-association-the-acm/> (дата обращения: 24.07.2024).
15. Hallamaa J., Kalliokoski T. AI Ethics as Applied Ethics // *Front. Comput. Sci.* 2022. Vol. 4. P. 776837. <https://doi.org/10.3389/fcomp.2022.776837>
16. Bentham J. An introduction to the principles of morals and legislation. Oxford: Clarendon Press, 1907. URL: <https://www.econlib.org/library/Bentham/bnthPML.html> (дата обращения: 26.07.2024).
17. Rigotti C., Fosch-Villaronga E. Fairness, AI & recruitment // *Computer Law & Security Review*. 2024. Vol. 53. P. 105966. <https://doi.org/10.1016/j.clsr.2024.105966>
18. 2023 Базовые модели, такие как ChatGPT, через призму Рекомендации ЮНЕСКО по этике искусственного интеллекта. URL: https://ai.gov.ru/knowledgebase/mezhdunarodnyedokumenty-po-razvitiyuii/2023_bazovye_modeli_takie_kak_chatgpt_cherez_prizmu_rekomenda-cii_yunesko_po_etike_iskusstvennogo_intellekta_foundation_models_such_as_chatgpt_through_the_prism_of_the_unesco_recommendation_on_the_ethics_of_artificial_intelligence_yunesko/ (дата обращения: 26.07.2024).
19. Wang X., Chen M., Chen N. How artificial intelligence affects the labour force employment structure from the perspective of industrial structure optimisation // *Heliyon*. 2024. Vol. 10, N 5. P. e26686. <https://doi.org/10.1016/j.heliyon.2024.e26686>
20. Акьюлов Р. И., Сквпень А. А. Роль искусственного интеллекта в трансформации современного рынка труда // *Дискуссия*. 2019. Вып. 94. С. 30–40.
21. Privacy and personal data risk governance for generative artificial intelligence: A Chinese perspective / Xiongbiao Ye, Yuhong Yan, Jia Li, Bo Jiang // *Telecomm Policy*. Central China Normal University. 2024. P. 102851. <https://doi.org/10.1016/j.telpol.2024.102851>
22. Chauhan C., Gullapalli R. R. Ethics of AI in Pathology // *Am. J. Pathol.* 2021. Vol. 191, N 10. P. 1673–1683. <https://doi.org/10.1016/j.ajpath.2021.06.011>
23. Renkema M., Tursunbayeva A. The future of work of academics in the age of Artificial Intelligence: State-of-the-art and a research roadmap // *Futures*. 2024. Vol. 163. P. 103453. <https://doi.org/10.1016/j.futures.2024.103453>
24. AI deception: A survey of examples, risks, and potential solutions / P. S. Park, S. Goldstein, A. O'Gara, M. Chen, Dan Hendrycks // *Patterns (NY)*. 2024. Vol. 5, N 5. P. 100988. <https://doi.org/10.1016/j.patter.2024.100988>
25. Кодекс этики в сфере искусственного интеллекта. Альянс в сфере искусственного интеллекта. URL: https://ethics.a-ai.ru/assets/ethics_files/2023/05/12/Кодекс_этики_20_10_1.pdf (дата обращения: 26.07.2024).
26. Этические аспекты нейротехнологий: ЮНЕСКО назначила международную группу экспертов для подготовки нового глобального стандарта. URL: <https://www.unesco.org/ru/articles/eticheskie-aspekty-neyrotekhnologiy-yunesko-naznachila-mezhdunarodnuyu-gruppu-ekspertov-dlya?hub=32618> (дата обращения: 20.07.2024).
27. OECD AI Principles overview. URL: <https://oecd.ai/en/ai-principles> (дата обращения: 28.07.2024).

28. On Artificial Intelligence-A European approach to excellence and trust White Paper on Artificial Intelligence A European approach to excellence and trust. European Commission. 2020. URL: https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en (дата обращения: 27.07.2024).
29. EU AI Act. European Union. URL: <https://artificialintelligenceact.eu/ai-act-explorer/> (дата обращения: 27.07.2024).
30. Filipova I. A. Legal Regulation of Artificial Intelligence: Experience of China // Journal of Digital Technologies and Law. 2024. Vol. 2, N 1. P. 46–73. <https://doi.org/10.21202/jdtl.2024.4>
31. Recommendation on the Ethics of Artificial Intelligence. URL: <https://unesdoc.unesco.org/ark:/48223/pf0000381137> (дата обращения: 26.09.2024).
32. Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence. URL: <https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/> (дата обращения: 28.07.2024).
33. The AIDA companion document: A roadmap for Canada's artificial intelligence law. URL: <https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act-aida-companion-document> (дата обращения: 27.07.2024).
34. Implementing the UK's AI regulatory principles: initial guidance for regulators. URL: <https://www.gov.uk/government/publications/implementing-the-uks-ai-regulatory-principles-initial-guidance-for-regulators> (дата обращения: 26.07.2024).
35. Global Index on Responsible AI 2024 / R. Adams, F. Adeleke, A. Florido, L. G. de Magalhães Santos, N. Grossman, L. Junck, K. Stone. URL: <https://ai.gov.ru/upload/iblock/2b0/1960deb403wb2ic179moquclt60i52gc.pdf> (дата обращения: 27.07.2024).

References

1. Novikov F.A. Symbolic artificial intelligence: mathematical foundations of knowledge representation. Moscow: Yurait; 2024. 278 p. (In Russ.)
2. Pospelov D.A. Modeling reasoning. Experience in analyzing mental acts. Moscow: Radio i svyaz'; 1989. 184 p. (In Russ.)
3. Stahl B.C., et al. Exploring ethics and human rights in artificial intelligence – A Delphi study. *Technol Forecast Soc Change*. 2023;191:122502. <https://doi.org/10.1016/j.techfore.2023.122502>
4. Romanov R.V. Classical concepts of solving artificial intelligence problems (philosophical aspect). *Prepodavatel' XXI vek = Teacher of the XXI Century*. 2012;(3):241–247. (In Russ.)
5. Kanysheva O.A. The teachings of Rene Descartes on the soul and consciousness. *Vestnik Voronezhskogo gosudarstvennogo universiteta. Seriya: Filosofiya = Bulletin of the Voronezh State University. Series: Philosophy*. 2021;(3):99-107. (In Russ.)
6. McCarthy J., et al. A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence. *AI Magazine*. 2006;(27).
7. Rai A., Sarker S. Editor's Comments Next-Generation Digital Platforms: Toward Human-AI Hybrids. *MIS Quarterly*. 2019;43(1):1–7.
8. Sinelnikova A.P. Will artificial intelligence be able to replace humans? *Vestnik nauki = Bulletin of Science*. 2020;(2):77-79. (In Russ.)
9. Barcaui A., Monat A. Who is better in project planning? Generative artificial intelligence or project managers? *Project Leadership and Society*. 2023;4:100101. <https://doi.org/10.1016/j.plas.2023.100101>

10. Korzynski P., et al. Generative artificial intelligence as a new context for management theories: analysis of ChatGPT. *Central European Management Journal*. 2023;31(1):0091. <https://doi.org/10.1108/CEMJ-02-2023-0091>
11. Nieto-Rodriguez A., Vargas R. How AI Will Transform Project Management. *PM World Journal*. 2023. P. 1–2. Available at: <https://pmworldlibrary.net/wp-content/uploads/2023/02/230224-hbr-how-ai-will-transform-project-management-publishing-news.pdf> (accessed 27.07.2024).
12. Mantelero A., Esposito M.S. An evidence-based methodology for human rights impact assessment (HRIA) in the development of AI data-intensive systems. *Computer Law & Security Review*. 2021;41:105561. <https://doi.org/10.1016/j.clsr.2021.105561>
13. 2024 Regulation of the European Union on Artificial Intelligence, ANO "Digital Economy". Available at: https://ai.gov.ru/knowledgebase/dokumenty-po-razvitiyu-ii-v-drugikh-stranakh/2024_reglament_evropeyskogo_soyuza_ob_iskusstvennom_intellekte_ano_cifrovaya_e_konomika/ (accessed 24.07.2024).
14. Lance B. Eliot. Responsible AI Relishes Preeminent Boost Via AI Ethics Proclamation By Top Professional Society The ACM. *Forbes*. Available at: <https://www.forbes.com/sites/lanceeliot/2022/11/27/responsible-ai-relishes-mighty-boost-via-ai-ethics-proclamation-rolled-out-by-esteemed-computing-profession-association-the-acm/> (accessed 24.07.2024).
15. Hallamaa J., Kalliokoski T. AI Ethics as Applied Ethics. *Front. Comput. Sci*. 2022;4:776837. <https://doi.org/10.3389/fcomp.2022.776837>
16. Bentham J. An introduction to the principles of morals and legislation. Oxford: Clarendon Press; 1907. Available at: <https://www.econlib.org/library/Bentham/bnthPML.html> (accessed 26.07.2024).
17. Rigotti C., Fosch-Villaronga E. Fairness, AI & recruitment. *Computer Law & Security Review*. 2024;53:105966. <https://doi.org/10.1016/j.clsr.2024.105966>
18. 2023 Basic models such as ChatGPT, through the prism of the UNESCO Recommendation on the Ethics of Artificial Intelligence. Available at: https://ai.gov.ru/knowledgebase/mezhdunarodnye-dokumenty-po-razvitiyuii/2023_bazovye_modeli_takie_kak_chatgpt_cherez_prizmu_rekomenda-cii_yunesko_po_etike_iskusstvennogo_intellekta_foundation_models_such_as_chatgpt_through_the_prism_of_the_unesco_recommendation_on_the_ethics_of_artificial_intelligence_yunesko/ (accessed 26.07.2024).
19. Wang X., Chen M., Chen N. How artificial intelligence affects the labour force employment structure from the perspective of industrial structure optimization. *Heliyon*. 2024;10(5):e26686. <https://doi.org/10.1016/j.heliyon.2024.e26686>
20. Ak'yulov R.I., Skovpen' A.A. The role of artificial intelligence in the transformation of the modern labor market. *Diskussiya = Discussion*. 2019;(94):30–40. (In Russ.)
21. Xiongbiao Ye, Yuhong Yan, Jia Li, Bo Jiang. Privacy and personal data risk governance for generative artificial intelligence: A Chinese perspective. *Telecomm Policy. Central China Normal University*. 2024. P. 102851. <https://doi.org/10.1016/j.telpol.2024.102851>
22. Chauhan C., Gullapalli R. R. Ethics of AI in Pathology. *Am. J. Pathol*. 2021;191(10):1673–1683. <https://doi.org/10.1016/j.ajpath.2021.06.011>
23. Renkema M., Tursunbayeva A. The future of work of academics in the age of Artificial Intelligence: State-of-the-art and a research roadmap. *Futures*. 2024;163:103453. <https://doi.org/10.1016/j.futures.2024.103453>
24. Park P.S., Goldstein S., O'Gara A., Chen M., Dan Hendrycks. AI deception: A survey of examples, risks, and potential solutions. *Patterns (NY)*. 2024;5(5):100988. <https://doi.org/10.1016/j.patter.2024.100988>

25. The Code of Ethics in the field of artificial intelligence. Al'yans v sfere iskusstvennogo intellekta. (In Russ.) Available at: https://ethics.a-ai.ru/assets/ethics_files/2023/05/12/Кодекс_этики_20_10_1.pdf (accessed 26.07.2024).

26. Ethical aspects of neurotechnology: UNESCO has appointed an international group of experts to prepare a new global standard. (In Russ.) Available at: <https://www.unesco.org/ru/articles/eticheskie-aspekty-neyrotekhnologiy-yunesko-naznachila-mezhdunarodnuyu-gruppu-ekspertov-dlya?hub=32618> (accessed 20.07.2024).

27. OECD AI Principles overview. Available at: <https://oecd.ai/en/ai-principles> (accessed 28.07.2024).

28. On Artificial Intelligence-A European approach to excellence and trust White Paper on Artificial Intelligence A European approach to excellence and trust. European Commission. 2020. Available at: https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en (accessed 27.07.2024).

29. EU AI Act. European Union. Available at: <https://artificialintelligenceact.eu/ai-act-explorer/> (accessed 27.07.2024).

30. Filipova I.A. Legal Regulation of Artificial Intelligence: Experience of China. *Journal of Digital Technologies and Law*. 2024;2(1):46–73. <https://doi.org/10.21202/jdtl.2024.4>

31. Recommendation on the Ethics of Artificial Intelligence. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000381137> (accessed 26.09.2024).

32. Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence. Available at: <https://www.whitehouse.gov/briefing-room/statements-releases/2023/10/30/fact-sheet-president-biden-issues-executive-order-on-safe-secure-and-trustworthy-artificial-intelligence/> (accessed 28.07.2024).

33. The AIDA companion document: A roadmap for Canada's artificial intelligence law. Available at: <https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act-aida-companion-document> (accessed 27.07.2024).

34. Implementing the UK's AI regulatory principles: initial guidance for regulators. Available at: <https://www.gov.uk/government/publications/implementing-the-uks-ai-regulatory-principles-initial-guidance-for-regulators> (accessed 26.07.2024).

35. Adams R., Adeleke F., Florido A., de Magalhães Santos L.G., Grossman N., Junck L., Stone K. Global Index on Responsible AI 2024. Available at: <https://ai.gov.ru/upload/iblock/2b0/1960deb403wb2ic179moquclt60i52gc.pdf> (accessed 27.07.2024).

Информация об авторах / Information about the Authors

Мальцева Ирина Федоровна, кандидат экономических наук, доцент кафедры региональной экономики и менеджмента, Юго-Западный государственный университет, г. Курск, Российская Федерация, e-mail: irina-fedorovna@inbox.ru, Researcher ID: O-3179-2015, ORCID: 0000-0002-8396-6657

Сазонова Елизавета Сергеевна, магистрант кафедры философии и социологии, Юго-Западный государственный университет, г. Курск, Российская Федерация, e-mail: sazonova.lizaa@yandex.ru

Irina F. Maltseva, Candidate of Sciences (Economics), Associate Professor of the Department of Regional Economics and Management, Southwest State University, Kursk, Russian Federation, e-mail: irina-fedorovna@inbox.ru, Researcher ID: O-3179-2015, ORCID: 0000-0002-8396-6657

Elizaveta S. Sazonova, Undergraduate of the Department of Philosophy and Sociology, Southwest State University, Kursk, Russian Federation, e-mail: sazonova.lizaa@yandex.ru