

УДК 1.37

*Научная статья*

## **КЛЮЧЕВЫЕ ПРОБЛЕМЫ, РИСКИ И ОГРАНИЧЕНИЯ ПРИМЕНЕНИЯ ИИ В МЕДИЦИНЕ И ОБРАЗОВАНИИ**

**Е.В. Брызгалина<sup>1,2</sup>, А.Н. Гумарова<sup>1</sup>, Е.М. Шкомова<sup>1\*</sup>**

<sup>1</sup> Московский государственный университет имени М.В. Ломоносова, 119991, Ленинские горы, МГУ, учебно-научный корпус «Шуваловский», г. Москва, Россия

<sup>2</sup> Балтийский федеральный университет им. Иммануила Канта, 236016, ул. Александра Невского, д. 14, г. Калининград, Россия

*Аннотация:* Целью статьи является выявление основных проблем, рисков и ограничений применения систем ИИ в социальной сфере (на примере образования и медицины), выделенных в научно-философских публикациях 2019 — начале 2022 г. К общим проблемам применения технологий ИИ в социальной сфере авторы статьи отнесли следующие: проблему достижимости доверенного ИИ, основанного на принципах доверенности, прозрачности, конфиденциальности и безопасности; проблему конфиденциальности данных; появление дискриминации и углубление неравенства из-за внедрения ИИ; трансформацию процессов взаимодействия субъектов при появлении «посредника» в виде технологии ИИ; необходимость разработки и внедрения правовых и этических ограничений взаимодействия пользователей с системами ИИ; риски принятия управленческих решений на основании данных ИИ. Описаны специфические проблемы, риски и ограничения внедрения ИИ в медицину и образование. Внедрение систем ИИ в социальную сферу требует создания комплексной системы этико-правового регулирования, предусматривающего формирование как общих, так и отраслевых правовых норм и функционирование социально-гуманитарной экспертизы как исследовательской деятельности по сопровождению проектов.

*Ключевые слова:* искусственный интеллект, этика искусственного интеллекта, искусственный интеллект в образовании; искусственный интеллект в медицине; социально-гуманитарная экспертиза

*Благодарности/Финансирование*

Данная работа выполнена в рамках проекта «Приоритет 2030».

---

\*© Брызгалина Е.В., Гумарова А.Н., Шкомова Е.М., 2022

## PHILOSOPHY OF EDUCATION

### KEY PROBLEMS, RISKS AND RESTRICTIONS OF USING ARTIFICIAL INTELLIGENCE IN MEDICINE AND EDUCATION

**E.V. Bryzgalina<sup>1,2</sup>, A.N. Gumarova<sup>1</sup>, E.M. Shkomova<sup>1</sup>**

<sup>1</sup> Lomonosov Moscow State University, Leninskie Gory, Moscow, Teaching and Scientific Building “Shuvalovsky”, 119991, Russia

<sup>2</sup> Immanuel Kant Baltic Federal University, A. Nevskogo st., 14, Kaliningrad, 236016, Russia

*Abstract:* The aim of this article is to identify the main problems, risks and restrictions of the use of AI systems in social life (on the examples of education and medicine) mentioned in scientific and philosophical publications in 2019–2021 and early 2022. It was revealed that general problems of applying AI technologies in the social life include the following: the problem of the attainability of trustworthy AI based on the principles of trust, transparency, confidentiality and security; data privacy issue; the emergence of discrimination and the extension of inequality caused by AI; transformation of interaction between subjects caused by the appearance of AI technologies as an “intermediary” of communication; the need to improve digital literacy; the need to develop and implement legal and ethical restrictions on user interaction with AI systems; prospects for making managerial decisions based on AI data. The article considers specific problems, risks and restrictions of the introduction of AI in medicine and education. The problems and risks of developing and implementing AI systems in social life require the creation of a comprehensive system of ethical and legal regulation, which provides for both the formation of general and sectoral legal norms, and the functioning of social and humanitarian expertise as a research activity in support of projects.

*Keywords:* artificial intelligence, ethics of AI, educational AI, AI in medicine, social and humanitarian expertise

#### *Acknowledgments/Financial Support*

This work was performed within the context of the “Priority 2030” project.

#### **Введение**

Расширение применения технологий искусственного интеллекта (ИИ)<sup>1</sup> в отраслях социальной сферы (образование, здравоохранение) обуславливается тем, что основные тенденции развития данных

---

<sup>1</sup> В данной статье понятия «искусственный интеллект (ИИ)», «технологии ИИ», «системы ИИ» используются как синонимичные, под ними подразумевается программное решение (алгоритм), предполагающее способность работы с данными, которому приписываются характеристики интеллектуальности («разумности», «рациональности», «возможности мыслить как человек» в определенных обстоятельствах).

технологий поддерживаются их применением и их использование одновременно является проявлением значимых трендов изменений социальной сферы. Целями применения ИИ в образовательной сфере выступают индивидуализация и персонификация процесса образования в различных формах; цифровизация оценивания и прогнозирования образовательных результатов; создание цифровой системы управления образованием, технологическое обеспечение принятия решений в образовательной политике [1]. Цели использования ИИ в медицине как науке и в здравоохранении обусловлены становлением так называемой персонализированной медицины, когда акцент на предикцию и превенцию заболеваний требует работы с массивами разнородных данных. Цифровизация систем образования и здравоохранения рассматривается как путь улучшения качества оказания соответствующих услуг гражданам, повышения эффективности управленческих решений. Анализ соотношения провозглашаемых целей внедрения ИИ в социальные сферы и формирующихся практик, а также возникающих и потенциальных рисков составляет предмет социально-философского исследования. Интерес к социально-философской проблематике ИИ связан с тем, что в медицине и образовании как социально чувствительных сферах происходит массовое соприкосновение человека с ИИ. Возникающий спектр позиций человека — создателя технологии, пользователя системы, субъекта, принимающего управленческое решение, и так далее — требует обращения к следствиям разрастающегося взаимодействия человека с системами ИИ, включая анализ рисков опосредования социальных отношений техническими системами. Множественность ценностных ориентаций и задачи смягчения/преодоления ценностных конфликтов обуславливают внимание к обоснованию ограничений на полном цикле разработок и внедрения технологий ИИ. Взаимное поддержание, с одной стороны, тенденций развития образовательной и медицинской практик, их концептуального базиса, с другой — целей применения систем ИИ определяет интенсивный рост соответствующих бизнес-проектов. Коммерциализация исследований в области ИИ, базирующаяся преимущественно на этической модели прагматизма, создает дополнительные основания для актуализации концептуального анализа целей, проблем, рисков и ограничений технологий ИИ от разработки и до использования.

ИИ интенсифицирует исследования образования на основе цифровых данных, а в применении к медицине формирует новую предметную область — цифровую биоэтику, использующую цифровые технологии для дискриптивного описания морального ланд-

шафта, значимого для медико-биологической сферы, и фиксации изменения содержания принципов биоэтики под влиянием смены общественных настроений.

### **Искусственный интеллект в медицине**

ИИ является инструментом, обеспечивающим процесс персонализации медицины и цифровизации системы здравоохранения [2]. Сбор и обработка разнородных данных кардинально изменяет представления о норме и патологии, о здоровье и здоровом образе жизни, о причинах заболеваний и методах борьбы с ними, трансформирует практики коммуникации пациентов и медицинского персонала. Е.В. Брызгалина описывает основные тренды трансформации медицины как науки и отрасли здравоохранения в направлении 4П медицины, ориентация на достижение которой неразрывно связана с применением систем искусственного интеллекта (ИИ) [3].

На основе анализа научно-философских публикаций последних лет выделим проблемы, требующие, по мнению их авторов, решений для медицины и системы здравоохранения. Подключение систем ИИ требует детализации содержания и механизмов принятия модели информированного добровольного согласия (ИДС) на использование в работе с пациентом технологий ИИ. Р. Хасс обращает внимание на то, что перед подписанием ИДС пациентам необходимо разобраться в принципах работы ИИ и в том, какие преимущества и какие риски возникают в результате диагностики и лечения с использованием ИИ [4]. С точки зрения Т. Плоуг и С. Холм, пациент должен обладать правом на отказ от диагностики и лечения с помощью ИИ. При этом система здравоохранения должна обеспечить как ответственное применение систем ИИ для достижения значимых для пациента результатов, так и качественную помощь при отказе от использования технологий ИИ в отношении конкретных пациентов [5].

При принятии решений с использованием технологий ИИ в медицине неизбежна трансформация отношений между врачом и пациентом. С. Бек отмечает, что решения, принимаемые ИИ, часто являются труднопредсказуемыми и влекут за собой ранее неизвестные риски, что вызывает необходимость обсуждения, в каких ситуациях решения могут приниматься машинами, а в каких это приведет к серьезным негативным последствиям [6]. От врачей и от пациентов в рамках парадигмы персонализированной медицины требуется понимание изменения целевых ориентиров и методов их достижения на базе новых возможностей медицинской науки. Врачу необходимо работать с новыми технологиями, осознавая свою роль

и роль машины в принятии решений, значимых для пациента. Возникает надежда, что с применением систем ИИ у врача будет больше времени на выявление индивидуальных особенностей пациента [4].

Поддержка ИИ экспертных врачебных решений относительно диагноза и индивидуальных схем лечения ведет к пересмотру представлений о классификации нозологий, изменяет архитектуру фиксации заболеваемости. Акцент смещается на профилактику индивидуальных рисков и достижение максимально возможного качества жизни конкретного пациента. Д. Мерз и Дж. Хобнер отмечают, что основная цель, которой должен служить ИИ, — это жизнь пациентов. Качество всех медицинских услуг должно быть улучшено. Поскольку медицинские исследования, технологическая модернизация системы здравоохранения происходят в рамках экономической политики и глобальных изменений, они должны быть организованы с соблюдением этических норм [7].

В связи с разработкой и внедрением ИИ в систему здравоохранения необходимо выработать дополнительные стандарты защиты персональных данных и обеспечения конфиденциальности информации о пациенте, в том числе должны быть определены лица, которые будут нести ответственность в ситуации «утечки» информации. Возникает задача пересмотра биоэтического принципа конфиденциальности в сторону расширенной трактовки его содержания. Й.А. Скорбург и др. рассматривают проблему сохранения конфиденциальности данных пациента в ситуации крупномасштабного сбора данных, необходимых для обучения ИИ [8]. Т. Вейчерт подчеркивает, что при отсутствии эффективной анонимизации необходимо соблюдать принципы Общего регламента по защите данных (GDPR), среди которых: законность, справедливость, прозрачность, ограничение целью, минимизация данных, точность, ограничение хранения данных, целостность и конфиденциальность, подотчетность. Если же ИИ должен принимать автоматизированные решения, необходимы дополнительные гарантии для предотвращения дискриминации или ошибочных результатов [9]. На это же обращают внимание Й. Скорбург и др., указывая на возможность появления новых форм дискриминации и углубления неравенства из-за внедрения ИИ в систему здравоохранения [8].

Очевидно, что проблемы и риски, обусловленные применением ИИ в медицине, требуют биоэтической оценки и являются вызовом для современной этики. В статье А. Манзесчке, У. Ниендерлаг говорится о бессмысленности ситуации, когда этическая оценка происходит на этапе внедрения инновационных продуктов, так как исправить что-либо уже невозможно [10]. Ф. Джоттерэнд и С. Боско, рассма-

тривая проблему дегуманизации медицины в результате внедрения ИИ, приходят к выводу, что по мере того, как ИИ становится все более распространенным явлением в клинической медицине, этические рамки, поддерживающие ответственное внедрение таких технологий, должны быть пересмотрены. Полагается, что необходимо выявить и проанализировать антропологические последствия ИИ в клиническом контексте [11].

Ф. Урсин и др. поднимают вопрос о том, каким именно этическим требованиям должен соответствовать ИИ. На основании публикаций специалистов в области радиологии по вопросу этичности использования ИИ в медицине они анализируют «объяснимость»<sup>2</sup> — принцип, направленный на большую предсказуемость алгоритмов машинного обучения [12]. Авторы ставят вопрос о концептуальной связи принципа объяснимости с классическими биоэтическими принципами: является ли объяснимость самостоятельным принципом? Несмотря на то что принципы, сформулированные Т. Бичампом и Д. Чилдрессом<sup>3</sup>, признаны в качестве универсальных, утверждается, что в каждой специальности существуют особые обстоятельства, требующие расширенного набора принципов [13]. Авторы утверждают, что расширение числа принципов может привести к потере статуса универсальности каждого принципа [12]. Делается вывод о том, что пока свойства объяснимости охватываются хотя бы одним из четырех принципов биомедицинской этики, объяснимость не следует признавать самостоятельным принципом, относя содержание к принципам «не навреди» и «делай благо».

Работа Е.Г. Гребенщиковой и П.Д. Тищенко отмечает важность анализа вызовов применения ИИ в конкретных областях медицины, в частности в хирургии. Авторы выделяют три режима включения медицинских систем ИИ в структуры отношений врачей и пациентов. Первый режим: ИИ выступает в качестве инструмента, улучшающего деятельность рук и органов чувств хирурга. Второй режим: ИИ играет роль ассистента или консультанта оперирующего хирурга. Для третьего, пока гипотетического, режима характерно

---

<sup>2</sup> Авторы говорят о требовании к ИИ быть прозрачным, объяснимым, подотчетным, доверенным. Для всех участников должны быть понятны принципы, на основании которых ИИ принимает решение; алгоритм должен находиться под контролем человека и может корректироваться в соответствии с актуальными проблемами. Отметим, что указанные принципы определяются исследователями по-разному. Поскольку отличия в определениях принципов достаточно тонкие, в данной статье они употребляются в качестве синонимичных.

<sup>3</sup> Т. Бичамп и Д. Чилдресс выделили четыре биоэтических принципа: «не навреди», «делай благо», принцип уважения автономии пациента, принцип справедливости.

превращение хирургического робота в автономного оперирующего хирурга. Отмечается необходимость междисциплинарного обсуждения вызовов ИИ в медицине, в котором биоэтике отводится роль модератора [14].

В статье Гребенщиковой Е.Г. и А.Г. Чучалина показано, что опыт обсуждения и решения сложных проблем, а также совещательные механизмы биоэтики отвечают актуальным вызовам биотехнологии и будут востребованы в будущем [15]. В поисках механизма, способного снизить экзистенциальные риски при внедрении ИИ, предлагается обратиться к этике ответственности Ганса Йонаса, философско-антропологическим идеям Б.Г. Юдина о человеке, постоянно изменяющем собственные границы [16;17].

Таким образом, на основании проанализированных публикаций отмечены специфичные для системы здравоохранения проблемы, к которым прежде всего относится проблема разработки и принятия модели информированного согласия в контексте диагностики и лечения с использованием ИИ; трансформация содержания принципа конфиденциальности; проблема выработки новой модели взаимоотношений врача и пациента. В качестве ключевых рисков применения ИИ в медицине выделены риски утечки персональной информации о пациенте, ошибки при диагностике и лечении с помощью ИИ, опасности появления дискриминации и углубления неравенства из-за внедрения ИИ в систему здравоохранения, возможность дегуманизации медицины. Внедрение систем ИИ в медицину требует повышения цифровой грамотности всех пациентов и медицинских работников. Отмечено, что система ограничений, направленная на предотвращение рисков, должна формироваться комплексно при биоэтическом и правовом регулировании полного цикла применения ИИ в медицине.

### **Искусственный интеллект в образовании**

Авторы рассмотренных публикаций анализируют влияние ИИ на образование или обращают внимание на социальные последствия его внедрения в эту сферу. В публикациях отмечается, что последствия распространения технологий ИИ в сфере образования обсуждаются недостаточно, а этические ограничения находятся на этапе разработки.

Рассмотрим основные проблемы, которые обозначают авторы при анализе перспективных областей использования ИИ в образовании. Создание единой концептуальной рамки доверенного ИИ в образовании осложняется включенностью на протяжении всего цикла реализации продукта множества субъектов, имеющих раз-

личные представления о целях его внедрения. У. Холмс и др. отмечают, что для создания этичного ИИ в образовании необходимо обеспечение свободного обмена информацией между экспертами, представляющими образование, моделирование алгоритмов и сбор больших данных [18]. В связи с этим в качестве базового принципа доверенного ИИ указывается принцип прозрачности, который предполагает, что вся информация об исходных положениях работы ИИ, о процессах принятия решений и сами принятые решения должны быть доступны для обмена между разработчиками, заказчиками и пользователями. На примерах «пограничных случаев» К. Китто и С. Кнайт демонстрируют, что существующие этические концепции, которые основаны на принципах уважения к личности, пользы и справедливости, не способны обеспечить реальное руководство для технических разработчиков систем ИИ, поскольку приводят к противоречиям. «Пограничные случаи» описывают следующие проблемы (1) информированного согласия пользователей в различных моделях использования ИИ в образовании; (2) хранения и предоставления данных субъектов образования, которые дали согласие на различные условия сбережения и использования их образовательных данных, полученных в результате групповой работы; (3) точности при персонализации образования и вариативности выбора в образовании в течение всей жизни. Исследователи предлагают создать открытую базу данных «пограничных случаев» для того, чтобы технические разработчики систем могли узнавать о реальных этических рисках и более тесно взаимодействовать с теми, кто занимается политикой, правом и этикой [19]. М. Чаудхри и др. называют прозрачность цифровых образовательных технологий ключевым принципом, создающим основу для других аспектов этичного ИИ, таких как объяснимость, интерпретируемость, справедливость, подотчетность и безопасность [20]. Поскольку прозрачность зависит от каждого актора, вовлеченного в разработку и использование ИИ, необходимо рассмотреть, как требования прозрачности различаются в группах заинтересованных сторон в образовательном контексте. Для этого авторы предлагают структуру индекса прозрачности, разработанную совместно с различными стейкхолдерами ИИ в образовании, включая преподавателей, экспертов в области образовательных технологий и специалистов-практиков ИИ. С. Винсент-Ланкрини, Рейер ван дер Влиес подчеркивают, что без осмысленного пользователя ИИ появляется риск злоупотребления выводами электронной системы [21].

Перспективы создания ИИ, ориентированного на человека, открывают возможности для разработки методик образования,

основанных на измерениях, и персонализации образовательных траекторий [22]. С одной стороны, персонализация образования требует сбора большого количества данных об учащихся, с другой, необходимо обеспечить сохранность данных. Проблема конфиденциальности, охраны от несанкционированного доступа при хранении и передаче данных пользователей актуальна в связи с высокой чувствительностью информации об индивидуальных особенностях, зафиксированных в образовательном процессе, особенно в отношении несовершеннолетних [23]. Защита данных пользователей в сфере образования обсуждается особенно остро, поскольку в образовании использование и обмен информацией не являются чисто коммерческой целью, тем не менее собранные данные представляют коммерческую ценность для заинтересованных компаний.

Ожидается, что ИИ в первую очередь будет способствовать демократизации образования — расширению доступа для различных социальных групп, в том числе для людей с особыми потребностями, к образовательному процессу. Между тем использование ИИ актуализирует риски монополизации данных в руках государства или компаний. Р. Паммер-Шиндлер выделяет этические ограничения ИИ в образовании в течение всей жизни. В неформальном образовании релевантные данные о профессиональном обучении относятся не только к учащимся, но также к их коллегам и клиентам. Отслеживание процесса обучения предполагает детализированный контроль учащихся как над личными данными, так и над групповыми [24].

В качестве отдельной проблемной области выделим трансформацию взаимодействия субъектов образования под влиянием ИИ. Появляются гибридные формы образования, в которые включены виртуальные помощники. Трансформируется роль учителя: приобретает особую ценность разумное применение педагогом разнообразных стратегий и методов, которые предлагает система ИИ. Л. Кобис и К. Мехнер предлагают переосмысление наставнической этики в ситуации взаимодействия пользователя с неживой системой при сохранении самого принципа наставничества, предполагающего высокую степень доверия и эмоциональной вовлеченности. Такие принципы, как благодеяние, непричинение вреда, справедливость, лояльность и забота, должны быть перенесены из сферы межличностных отношений в отношения пользователя и цифрового помощника. В то время как традиционное человеческое наставничество носит личный характер, а его этические акценты сосредоточены на межличностных отношениях, принципы этики ИИ сосредоточены на использовании технологии человеком и ее значении для общества.

В отличие от этики наставничества этика ИИ концентрируется на аспектах «технической надежности», «безопасности», «демократичности» и др. [25]. Д. Ричардз и В. Дигнум говорят о необходимости разработки подхода «ценностного дизайна» для создания образовательных чат-ботов: в основу алгоритмов должны быть заложены человеческие ценности и общепринятые социальные нормы [26].

Следствием недостижения прозрачного ИИ в образовании может стать предвзятость систем ИИ и образовательная дискриминация. Е. Руан и др. фиксируют необходимость использовать множественные подходы при разработке чат-бота для конкретных сценариев вместо применения стандартных абстрактных принципов. Отсутствие вариативности в содержании образовательных платформ и в оценке результатов образования может привести к подчинению всех пользователей узкому образцу учащегося определенного «идеального» типа [27]. При обсуждении проблемы алгоритмического предубеждения ставятся вопросы об ограниченности вводной выборки данных для обучения алгоритма, а также о ценностных основаниях деятельности специалистов-разработчиков системы.

С появлением возможности анализа больших массивов данных у теоретиков образования растет надежда на построение так называемого «точного образования». Б. Уилльямсон описывает новое направление в когнитивистике, в котором при помощи систем ИИ анализируются биометрические данные обучающегося [28]. Беспокойство вызывает то, что влияние ИИ на образование выходит за рамки отслеживания деятельности учащихся на обучающих платформах, достигая соматических и психологических субстратов обучения. При этом результаты подобного исследования образования во многом основаны на артефактах, которые сконструированы самими же вычислительными системами, вследствие чего происходит отдаление от объекта познания. С.Б. Шам и Р. Лакин отмечают риски того, что при исследовании образования через количественные показатели будут утеряны специфические, присущие образованию как особой человеческой деятельности неизмеримые характеристики [29].

Таким образом, в профильных публикациях выделяются следующие проблемы внедрения ИИ в образование: проблема достижения доверенного ИИ, при создании и использовании которого технические разработчики, заказчики образовательных продуктов, специалисты в теории и практике образования и пользователи находятся в коммуникации относительно целей, перспектив и социогуманитарных рисков использования ИИ в образовании. Специ-

фика рассмотрения проблемы конфиденциальности связана с тем, что зачастую личные данные, необходимые для персонификации образования, оказываются объектами коммерческого использования. Отличающей применением ИИ в сфере образования является и проблема хранения/использования групповой информации. Трансформация взаимодействий субъектов образования ставит задачи создания систем, чувствительных к вариативным образовательным запросам пользователей и работающих на основании принципов этико-правового взаимодействия. Доверие к системам ИИ как источнику достоверного знания об образовании является условием датафикации управленческих решений в сфере образования. В качестве основных авторами выделяются риски, связанные с экзистенциальной безопасностью человека и антропологической сущностью образования.

В данной статье на основании анализа публикаций было выявлено, что к общим проблемам применения технологий ИИ в социальной сфере относят следующие: проблему достижимости этического ИИ, основанного на принципах доверенности, прозрачности, конфиденциальности и безопасности, что весьма значимо в сфере медицины и образования; проблему конфиденциальности чувствительных медицинских и образовательных данных; появление новых форм дискриминации и углубления неравенства из-за внедрения ИИ; трансформацию процессов взаимодействия субъектов в медицине и образовании; необходимость разработки и внедрения правовых и этических ограничений взаимодействия пользователей с системами ИИ; перспективы принятия управленческих решений в области здравоохранения и образования на основании данных. Среди важных задач, которые необходимо решить для развития технологий ИИ для медицины и образования, следует отметить создание инфраструктуры для коммуникации и взаимодействия технических специалистов, теоретиков и практиков, различных групп пользователей (например, в образовании — от родителей и учителей до детей и взрослых учащихся, в медицине — от врачей до членов семей пациентов).

Специфичными для сферы здравоохранения авторы считают проблемы, связанные с разработкой и принятием модели ИДС на использование в работе с пациентом технологий ИИ, пересмотр содержания биоэтического принципа конфиденциальности, распределение ответственности для разных уровней использования врачом ИИ как системы поддержки врачебных решений.

Применительно к сфере образования можно сделать вывод о том, что специфичными являются следующие проблемы: проблема

конфиденциальности и сочетания личной и групповой информации пользователей образовательных сред на основе ИИ, в том числе полученной относительно несовершеннолетних граждан, не способных дать юридически значимое согласие и в отношении которых возникает вопрос о допустимости «открытого будущего», связанного с отказом от цифровых образовательных следов; отсутствие ценностных и правовых оснований использования конкретных технологий (например, образовательных и кураторских чат-ботов); трансформация взаимодействия субъектов образования; построение управления образованием на основании цифровых данных.

Необходимость регулирования полного цикла разработки и применения систем ИИ в социальной сфере обусловлена высокой рискогенностью практик использования ИИ в качестве систем поддержки экспертных решений. Среди рисков, присущих как сфере образования, так и медицины, в статье выделены следующие: риски для автономии субъектов, риски утечки данных, ошибок систем и злоупотреблений доверием системам ИИ. Применение ИИ в образовании сопряжено со следующими рисками: утечкой личных и групповых данных; утратой вариативности выбора образовательных траекторий; дискриминацией пользователей по когнитивным способностям или иным характеристикам; объективизацией учащихся как источников больших данных; потерей антропологического содержания образования и его алгоритмизацией. В медицине внедрение технологий ИИ должно учитывать особенности тесного переплетения медицинской науки и практики, широкие возможности вмешательства в физическую и психическую целостность человека. Специфичные для медицины риски использования систем ИИ связаны с возможной утечкой персональной информации о пациенте, ошибками при диагностике и лечении с помощью ИИ, появлением новых форм дискриминации и углубления неравенства из-за внедрения ИИ в систему здравоохранения, а также с риском дегуманизации медицины в целом. Дифференциация рисков и ограничений полного цикла ИИ в медицине и образовании — от разработки до использования — должна быть рассмотрена дополнительно.

В статье выделены как проблемы внутри образования и здравоохранения, имеющие потенциал локальных этико-правовых решений, так и проблемы, требующие социально-философского рассмотрения, поскольку они представляют собой вызовы для понимания антропологической сущности образования и онтологических смыслов врачевания. Решение проблем и минимизация рисков разработки и внедрения систем ИИ в социальную сферу зависит от

создания комплексной системы этико-правового регулирования, предусматривающего формирование как общих, так и отраслевых правовых норм и функционирование социально-гуманитарной экспертизы как исследовательской деятельности по сопровождению проектов.

## СПИСОК ИСТОЧНИКОВ

1. *Брызгалова Е.В.* Искусственный интеллект в образовании: Анализ целей внедрения // *Человек*. 2021. Т. 32, № 2. С. 9–29.
2. *Aach M., Plattner H.* Die Zukunft der Medizin. 2019. 428 s.
3. *Брызгалова Е.В.* Медицина в оптике искусственного интеллекта: философский контекст будущего // *Человек*. 2019. Т. 30, № 6. С. 54–71.
4. *Huss R.* Künstliche Intelligenz, Robotik und Big Data in der Medizin. Berlin; Heidelberg: Springer, 2019. 124 s.
5. *Ploug T., Holm S.* The right to refuse diagnostics and treatment planning by artificial intelligence // *Medicine, Health Care and Philosophy*. 2020. Vol. 23, N 1. P. 107–114.
6. *Beck S.* Künstliche Intelligenz–ethische und rechtliche Herausforderungen // *Philosophisches Handbuch Künstliche Intelligenz*. 2020. S. 1–28.
7. *Merz D., Hübner J.* Bessere Medizin? Künstliche Intelligenz verantwortlich gestalten // *Künstliche Intelligenz im Gesundheitswesen*. Gabler; Wiesbaden: Springer, 2022. S. 253–283.
8. *Skorburg J.A., Sinnott-Armstrong W., Conitzer V.* Ai methods in bioethics // *AJOB Empirical Bioethics*. 2020. Vol. 11, N 1. P. 37–39.
9. *Weichert T.* Datenschutz im Kontext der medizinischen Nutzung von KI-Systemen: heute und morgen= Privacy in the context of medical use of AI systems: today and tomorrow // *Zeitschrift für medizinische Ethik*. 2021. Bd 67, N 3. S. 351–365.
10. *Manzeschke A., Niederlag W.* (ed.). Ethische Perspektiven auf Biomedizinische Technologie. Walter de Gruyter GmbH & Co KG, 2020. Bd 3. 341 s.
11. *Jotterand F., Bosco C.* Artificial intelligence in medicine: A sword of damocles? // *Journal of Medical Systems*. 2022. Vol. 46, N 1. P. 1–5.
12. *Ursin F., Timmermann C., Steger F.* Explicability of artificial intelligence in radiology: Is a fifth bioethical principle conceptually necessary? // *Bioethics*. 2022. Vol. 36, N 2. P. 143–153.
13. *Beauchamp T.L., Childress J.F.* Principles of biomedical ethics. 4th ed. Oxford: Oxford University Press, 1994. 546 p.
14. *Гребеницкова Е.Г., Тищенко П.Д.* Биоэтические вызовы искусственного интеллекта в хирургии // *Клиническая и экспериментальная хирургия*. 2021. Т. 9, №. 3. С. 7–15.
15. *Гребеницкова Е.Г., Чучалин А.Г.* Биоэтика в XXI веке: вызовы и перспективы // *Медицинская этика*. 2021. С. 10–15.
16. *Бадмаева М.Х.* Этика искусственного интеллекта: принцип ответственности Ганса Йонаса // *Вестник Бурятского государственного университета. Философия*. 2022. № 1. С. 67–79.
17. *Гребеницкова Е.Г., Тищенко П.Д.* Оцифрованное будущее медицины: вызовы для биоэтики // *Философские науки*. 2020. Т. 63, № 2. С. 83–103.
18. *Holmes W., Porayska-Pomsta K., Holstein K., Sutherland E., Baker T., Shum S.B., Santos O.C., Rodrigo M.T., Cukurova M., Bittencourt I.I., Koedinger K.R.* Ethics of AI in education: Towards a community-wide framework // *International Journal of Artificial Intelligence in Education*. 2021. 25 p.

19. Kitto K., Knight S. Practical ethics for building learning analytics // British Journal of Educational Technology. 2019. Vol. 50, is. 6. P. 2779–3455.
20. Chaudhry M., Cukurova M., Luckin R. A transparency Index framework for AI in education. 2022. 17 p.
21. Vincent-Lancrin S., Vlies R. van der. Trustworthy artificial intelligence (AI) in education: Promises and challenges // OECD Education Working Papers. 2020. N 218. 15 p.
22. Renz A., Vladova G. Reinvigorating the discourse on Human-Centered Artificial Intelligence in educational technologies // Technology Innovation Management Review. 2021. Vol. 11(5). P. 5–16.
23. Chaudhry M., Kazim E. Artificial Intelligence in education (AIEd): a high-level academic and industry note 2021 // AI Ethics. 2022. N 2. P. 157–165.
24. Pammer-Schindler V., Rosé C. Data-related ethics issues in technologies for informal professional learning // International Journal of Artificial Intelligence in Education. 2021. P. 609–635.
25. Köbis L., Mehner C. Ethical questions raised by AI-supported mentoring in higher education // Frontiers of Artificial Intelligence. 2021. 43 p.
26. Richards D., Dignum V. Supporting and challenging learners through pedagogical agents: Addressing ethical issues through designing for values // British Journal of Educational Technologies. 2019. N 50. P. 2885–2901.
27. Ruane E., Birhane A., Ventresque A. Conversational AI: Social and ethical considerations. 2019. 12 p.
28. Williamson B. New digital laboratories of experimental knowledge production: Artificial intelligence and education research // London Review of Education. 2020. Vol. 18(2). P. 209–220.
29. Shum B.S.J., Luckin R. Learning analytics and AI: Politics, pedagogy and practices // British Journal of Educational Technologies. 2019. Vol. 50. P. 2785–2793.

## REFERENCES

1. Bryzgalina E.V. Artificial Intelligence in Education. Analysis of Implementation Goals. *Chelovek*. 2021. Vol. 32. N 2. P. 9–29. (In Russ.)
2. Aach M., Plattner H. Die Zukunft der Medizin. 2019. 428 s.
3. Bryzgalina E.V. Medicine in the optics of artificial intelligence: the philosophical context of the future. *Chelovek*. 2019. T. 30, N 6. P. 54–71. (In Russ.)
4. Huss R. Künstliche Intelligenz, Robotik und Big Data in der Medizin. Springer Berlin Heidelberg, 2019. 124 s.
5. Ploug T., Holm S. The right to refuse diagnostics and treatment planning by artificial intelligence. *Medicine, Health Care and Philosophy*. 2020. T. 23, N 1. P. 107–114.
6. Beck S. Künstliche Intelligenz – ethische und rechtliche Herausforderungen. In: *Philosophisches Handbuch Künstliche Intelligenz*. 2020. S. 1–28.
7. Merz D., Hübner J. Bessere Medizin? Künstliche Intelligenz verantwortlich gestalten. In: *Künstliche Intelligenz im Gesundheitswesen*. Springer Gabler, Wiesbaden, 2022. S. 253–283.
8. Skorburg J.A., Sinnott-Armstrong W., Conitzer V. Ai methods in bioethics. *AJOB Empirical Bioethics*. 2020. T. 11, N 1. P. 37–39.
9. Weichert T. Datenschutz im Kontext der medizinischen Nutzung von KI-Systemen: heute und morgen. *Zeitschrift für medizinische Ethik*. 2021. T. 67, N 3. S. 351–365.
10. Manzeschke A., Niederlag W. (ed.). Ethische Perspektiven auf Biomedizinische Technologie. Walter de Gruyter GmbH & Co KG, 2020. T. 3.

11. Jotterand F., Bosco C. Artificial Intelligence in Medicine: A Sword of Damocles? *Journal of Medical Systems*. 2022. T. 46, N 1. P. 1–5.
12. Ursin F., Timmermann C., Steger F. Explicability of artificial intelligence in radiology: Is a fifth bioethical principle conceptually necessary? *Bioethics*. 2022. T. 36. N 2. P. 143–153.
13. Beauchamp T.L., Childress J.F. Principles of biomedical ethics. 4th ed. Oxford: Oxford University Press, 1994. 546 p.
14. Grebenshchikova E.G., Tishchenko P.D. Bioethical challenges of artificial intelligence in surgery // *Clinical and experimental surgery*. 2021. T. 9, N 3. P. 7–15. (In Russ.)
15. Grebenshchikova E.G., Chuchalin A.G. Bioethics in the 21st century: challenges and perspectives. *In: Medical ethics*. 2021. P. 10–15. (In Russ.)
16. Badamaeva M.Kh. Ethics of artificial intelligence: Hans Jonas' principle of responsibility. *The Buryat State University Bulletin. Philosophy*. 2022. N 1. P. 67–79. (In Russ.)
17. Grebenshchikova E.G., Tishchenko P.D. Digitized Future of Medicine: Challenges for Bioethics. *Philosophical Science*. 2020. T. 63. N 2. P. 83–103. (In Russ.)
18. Holmes W., Porayska-Pomsta K., Holstein K., Sutherland E., Baker T., Shum S.B., Santos O.C., Rodrigo M. T., Cukurova M., Bittencourt I.I., Koedinger K.R. Ethics of AI in Education: Towards a Community-Wide Framework. *International Journal of Artificial Intelligence in Education*. 2021. 25 p.
19. Kitto K., Knight S. Practical ethics for building learning analytics. *British Journal of Educational Technology*. 2019. Vol. 50, is. 6. P. 2779–3455.
20. Chaudhry M., Cukurova M., Luckin R. A Transparency Index Framework for AI in Education. 2022. 17 p.
21. Vincent-Lancrin S., van der Vlies R. Trustworthy artificial intelligence (AI) in education: Promises and challenges. *In: OECD Education Working Papers*. 2020. N 218. 15 p.
22. Renz A., Vladova G. Reinvigorating the Discourse on Human-Centered Artificial Intelligence in Educational Technologies. *Technology Innovation Management Review*. 2021. 11(5). P. 5–16.
23. Chaudhry M., Kazim E. Artificial Intelligence in Education (AIEd): a high-level academic and industry note 2021. *In: AI Ethics*. 2022. 2. P. 157–165.
24. Pammer-Schindler V., Rosé C. Data-Related Ethics Issues in Technologies for Informal Professional Learning. *International Journal of Artificial Intelligence in Education*. 2021. P. 609–635.
25. Köbis L, Mehner C. Ethical Questions Raised by AI-Supported Mentoring in Higher Education. *In: Frontiers of Artificial Intelligence*. 2021. 43 p.
26. Richards D., Dignum V. Supporting and challenging learners through pedagogical agents: Addressing ethical issues through designing for values. *British Journal of Educational Technologies*. 2019. N 50. P. 2885–2901.
27. Ruane E., Birhane A., Ventresque A. Conversational AI: Social and Ethical Considerations. 2019. 12 p.
28. Williamson B. New digital laboratories of experimental knowledge production: Artificial intelligence and education research. *London Review of Education*. 2020. 18(2). P. 209–220.
29. Shum B.S.J., Luckin R. Learning analytics and AI: Politics, pedagogy and practices. *British Journal of Educational Technologies*. 2019. 50. P. 2785–2793.

**Информация об авторах:** *Брызгалина Елена Владимировна* — кандидат философских наук, доцент, заведующий кафедрой философии образования философского факультета МГУ имени М.В. Ломоносова; Балтийский федеральный университет им. Иммануила Канта, тел.: +7 (495) 939-20-08; [evbrz@yandex.ru](mailto:evbrz@yandex.ru)

*Гумарова Анастасия Николаевна* — аспирант кафедры философии образования философского факультета МГУ имени М.В. Ломоносова, +7 (495) 939-20-08; [anastasiya-gumarova@yandex.ru](mailto:anastasiya-gumarova@yandex.ru)

*Шкомова Екатерина Михайловна* — кандидат философских наук, ассистент кафедры этики философского факультета МГУ имени М.В. Ломоносова, тел.: +7 (494) 939-20-55; [eshkomova@yandex.ru](mailto:eshkomova@yandex.ru)

**Information about the authors:** *Bryzgalina Elena Vladimirovna* — Candidate of Philosophical Science, Associate Professor, Department of Philosophy of Education, Faculty of Philosophy, Lomonosov Moscow State University, Immanuel Kant Baltic Federal University, tel.: +7 (495) 939-20-08; [evbrz@yandex.ru](mailto:evbrz@yandex.ru)

*Gumarova Anastasiya Nikolaevna* — postgraduate student, Department of Philosophy of Education, Faculty of Philosophy, Lomonosov Moscow State University, tel.: +7 (495) 939-20-08; [anastasiya-gumarova@yandex.ru](mailto:anastasiya-gumarova@yandex.ru)

*Shkomova Ekaterina Mikhailovna* — Candidate of Philosophical Science, assistant, Department of Ethics, Faculty of Philosophy, Lomonosov Moscow State University, tel.: +7 (494) 939-20-55; [eshkomova@yandex.ru](mailto:eshkomova@yandex.ru)

Поступила в редакцию 05.06.2022;  
принята к публикации 09.09.2022