

ISSN 2524-2369 (Print)

ISSN 2524-2377 (Online)

УДК [612.089:004.8]:608.1

<https://doi.org/10.29235/2524-2369-2024-69-2-95-107>

Поступила в редакцию 27.02.2023

Received 27.02.2023

ФІЛАСОФІЯ І САЦЫЯЛОГІЯ
PHILOSOPHY AND SOCIOLOGY

В. Н. Сокольчик¹, А. И. Разуванов²

¹*Институт философии Национальной академии наук Беларуси, Минск, Беларусь*

²*Республиканский научно-практический центр медицинской экспертизы и реабилитации, Минск, Беларусь*

**ЭТИЧЕСКИЕ ПРИНЦИПЫ ИСПОЛЬЗОВАНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА
В ИССЛЕДОВАТЕЛЬСКОЙ ПРАКТИКЕ
(НА ПРИМЕРЕ БИОМЕДИЦИНСКИХ ИССЛЕДОВАНИЙ)**

Аннотация. Рассматривается актуальная для современной науки тема – этическое сопровождение исследований, выполняемых с применением искусственного интеллекта (ИИ). Несмотря на значительное количество зарубежных и отечественных публикаций, посвященных теме ИИ, концептуальное обоснование этики применения ИИ в научных исследованиях остается неразработанным. Основываясь на международных рекомендациях и статьях, а также на собственном опыте исследовательской деятельности и работы в этических комитетах, авторы определяют и анализируют базовые этические принципы использования ИИ в научных исследованиях. Предложенные принципы рассматриваются в контексте их практического применения в сфере биомедицины для защиты интересов человека и природы, в том числе для сохранения конфиденциальности данных участников, препятствия дискриминации, защиты от ошибок ИИ, уважения информированного согласия, а также соблюдения норм «открытой науки», взаимного доверия разработчиков и пользователей и др. Применение предложенных принципов ориентирует ученых, разработчиков ИИ, этические комитеты, проводящие экспертизу исследований, общество в целом на приоритеты гуманизации науки, уважение человека и природы, а также на просвещение общества, создание нормативно-правовой базы, этических рекомендаций и кодексов этики применения ИИ в научной деятельности.

Ключевые слова: искусственный интеллект (ИИ), открытая наука, научное исследование, этика, управляемость, безопасность, объяснимость, эффективность, справедливость, доверие, этические комитеты

Для цитирования: Сокольчик, В. Н. Этические принципы использования искусственного интеллекта в исследовательской практике (на примере биомедицинских исследований) / В. Н. Сокольчик, А. И. Разуванов // Вест. Нац. акад. наук Беларусі. Сер. гуманітар. навук. – 2024. – Т. 69, № 2. – С. 95–107. <https://doi.org/10.29235/2524-2369-2024-69-2-95-107>

Valerya N. Sokolchik¹, Aliaksei I. Razuvanau²

¹*Institute of Philosophy of the National Academy of Sciences of Belarus, Minsk, Belarus*

²*National Science and Practice Centre of Medical Assessment and Rehabilitation, Minsk, Belarus*

**ETHICAL PRINCIPLES FOR USING THE ARTIFICIAL INTELLIGENCE IN RESEARCH
(BASED ON BIOMEDICAL RESEARCH)**

Abstract. The article is devoted to the relevant issue of modern science and ethical support for scientific research using Artificial Intelligence (AI). Despite a significant number of foreign and domestic publications about AI, the conceptual framework for the ethics of scientific research using AI remains undeveloped. Based on the international recommendations and articles, as well as own research experience and membership of research ethical committees, the authors define and analyze the basic ethical principles for the scientific research using AI. The proposed principles are considered in the context of their practical application in the field of biomedicine, which are connected with protection of the mankind and nature, maintaining the confidentiality of participants' data, preventing discrimination, protecting against errors of AI, respecting informed consent, as well as observing the norms of “open science”, mutual trust from developers and users, etc. The application of the

proposed principles orient the scientists, the developers of artificial intelligence, ethical committees, realizing review process, all society, to the priority of the humanization of science, respect for man and nature, as well as education of society regarding to AI, the creation of a regulatory framework, ethical recommendations and codes of ethics for the using of AI in scientific research.

Keywords: Artificial Intelligence (AI), Open Science, scientific research, ethics, governability, safety, explainability, efficiency, justice, trust, ethics committees

For citation: Sokolchik V. N., Razuvanau A. I. Ethical principles for using the artificial intelligence in research (based on biomedical research). *Vestsi Natsyonal'noi akademii navuk Belarusi. Seriya humanitarnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Humanitarian Series*, 2024, vol. 69, no. 2, pp. 95–107 (in Russian). <https://doi.org/10.29235/2524-2369-2024-69-2-95-107>

Введение. Использование искусственного интеллекта (ИИ) сегодня является основой практически любой деятельности. Современная наука активно внедряет системы ИИ, но особенно актуальным это становится для новой парадигмы «open science» (открытая наука), существование которой, по сути, определяется взаимодействием с ИИ. Открытая наука основывается на машинном интеллекте как в методологии и используемых технологиях, так и при анализе «больших данных», обработке полученных результатов, а также при распространении данных исследований через открытые источники (такие, как научные платформы, социальные сети и т. д.) [1].

Применение ИИ в науке, особенно связанной с изучением человека, общества, биосферы, априори требует гуманистического сопровождения исследований, предполагая формирование этико-правовых установок, определяющих границы и рамки использования ИИ. Независимо от нашего восприятия ИИ в его «слабом» (как машин и алгоритмов, решающих конкретные задачи) или «сильном» варианте (как искусственного аналога разума, способного к мультифункциональности и самообучению) этические принципы построения и взаимоотношения с ИИ становятся необходимой основой для технологий ИИ, выстраивая рамки (само)ограничения ИИ, направленного на сохранение человеческого и природного существования, улучшение качества жизни человека.

Ни одно современное научное исследование, в частности биомедицинское, не может осуществляться вне следования этическим нормам, поскольку предполагает использование персональных данных человека, вмешательство в его личную жизнь, манипуляции с его организмом, транслирование его мнений и т. д. Инструментами этической коррекции научных исследований являются, во-первых, этическая грамотность и подготовленность самого ученого и исследовательской команды, во-вторых, исследовательские этические комитеты (в медицине – независимые этические комитеты (НЭКи)), осуществляющие этическую экспертизу научных проектов, в-третьих, собственно общественность, представленная организациями и личностями, оказывающими влияние на постановку вопросов, дизайн исследования, распространение его данных и т. д.

Сфера исследовательской этики в целом урегулирована, включая документы (как международные, так и национальные), рекомендации, действующие этические стандарты и структуры (этические комитеты и комиссии) и т. д., но на практике о решении этических вопросов исследования задумываются далеко не всегда, зачастую предпочитая интересы и нужды науки интересам конкретных людей и природы. Соответственно, если общие принципы исследовательской этики, активно развивающейся с середины XX в., до сих пор не всегда рассматриваются в ряду приоритетов исследования, то вопросы этики работы с ИИ являются для современной науки новыми и неизученными вызовами.

Работы по формированию этики использования ИИ, в частности этики научных исследований с использованием ИИ, в современном мире ведутся очень активно. За последние 3–5 лет появилось достаточно большое количество документов, предлагающих этические регулятивы и рекомендации для разработчиков и пользователей систем ИИ. Среди наиболее значимых – рекомендации Совета по искусственному интеллекту организации экономического сотрудничества и развития (2020) [2]; этические рекомендации Международной коалиции регуляторных органов в области лекарственных средств (ICMRA) для сфер клинической медицины и фармацевтики (2021) [3]; заявление об искусственном интеллекте, робототехнике и «автономных» системах Совета Европы (2020) [4], рекомендации по этике использования искусственного интеллекта

ЮНЕСКО (2021) [5], руководство ВОЗ по этике и управлению использованием искусственного интеллекта в здравоохранении (2021) [6], Кодекс этики искусственного интеллекта Российской Федерации (2022) [7] и др. В стадии разработки находится европейский проект, реализующий построение платформы европейской открытой науки (EOSC), включающий разработку этических стандартов исследований с применением ИИ, а также рекомендации по проведению этической экспертизы исследований с применением ИИ (2022) европейского института науки Ады Лавлейс [8].

При всем кажущемся многообразии публикаций и документов по вопросам этики ИИ необходимо констатировать, что такие публикации преобладают в Западной Европе, Америке и России, в Беларуси же вопросы исследовательской этики с применением ИИ практически не рассматриваются. Методологической проблемой изучения этики ИИ становится отсутствие рефлексии над выявлением базовых принципов этики ИИ, обозначение которых позволило бы создать фундамент, на основе которого развиваются механизмы этического осмысления, обучения как человека, так и ИИ.

Цель статьи, подготовленной по результатам междисциплинарной инициативы по осмыслению этики ИИ, – проанализировать этическую систему использования ИИ в научной деятельности, выстроив иерархию взаимосвязи между этическими принципами работы с ИИ, возникающими при применении этих принципов, и способами их разрешения, актуальными для современной науки и общества.

Авторы при подготовке материала основывались на изучении современных документов, рекомендаций, кодексов, научных проектов и статей, учитывая интервью с разработчиками и пользователями ИИ в сфере научных исследований, а также базировались на собственном опыте научной деятельности в медицине и здравоохранении и опыте экспертиз этических комитетов системы здравоохранения.

Этические принципы использования ИИ в исследованиях. Этическими императивами использования ИИ в исследованиях, связанных с человеком и природой, являются главные принципы современной биоэтики. В них аккумулируются направленность всех осуществляемых действий на благо человека и природы («делай добро»), идея непричинения вреда живому («не навреди»), а также признание права человеческой личности сохранять свои ценности и определять границы своей самости (принцип автономии).

Однако для реализации этических императивов при создании систем ИИ и последующей работе с ними необходимо также разработать принципы, определяющие взаимодействие между ИИ и человеком и включающие необходимые требования к такому взаимодействию. Первый блок таких принципов – безопасность (для человека, общества, природы), управляемость, объяснимость, эффективность – условно может быть назван «объективным». Этот блок определяет надличностные этические требования, предъявляемые непосредственно к системам ИИ, этико-правовым установкам общества, а также критериям научности и обоснованности использования ИИ.

Второй блок принципов, включающий принцип справедливости и принцип доверия, условно может быть назван «субъективным», поскольку в большей степени определяется внутренними установками личности (общества), связанными с включенностью в процесс взаимодействия с ИИ и личностно окрашенным восприятием этого процесса.

Безусловно, предложенный список принципов можно расширять и дополнять, однако обозначенные конструкты являются базовыми, необходимыми как для разработчиков систем ИИ, так и для пользователей ИИ – исследователей, операторов и т. д., а также для НЭКов, проводящих экспертизу научных исследований [9]. Для разъяснения вышеназванных принципов необходимо остановиться на их функционале несколько подробнее.

Принцип безопасности. Принцип безопасности предполагает учитывать несколько параметров, необходимых для корректной работы с системами ИИ [10]. Прежде всего, безопасность содержит надежность и предсказуемость действий систем, включая четкое определение критериев надежности ИИ для человека, общества и природы, предустановленность системы анализа и исправления ошибок, продуманность системы защиты от рисков [11].

Говоря о надежности в сфере исследовательской этики, мы не рассматриваем исключительно техническую надежность системы. Здесь речь идет о разработке таких параметров надежности, которые определяют безопасность человека и природы. При этом наряду с физической безопасностью, предполагающей отсутствие физического вреда и дискомфорта для человека и природы, существует также психическая, социокультурная и другая безопасность человека. Она связана с уважением установок человеческой личности, бережным отношением к его ценностям и идеалам, неразглашением личных данных, уважением к принятому человеком решению и т. д.

Важное значение для исследований с применением ИИ имеет определение групповых и личностных рисков для участников исследований. Если групповые риски исследователь, как правило, рассматривает при продумывании дизайна исследования, то внимание к личностным рискам минимальное. Личностные риски могут возникать как по причине физического различия испытуемых (например, в медицине), так и в зависимости от ментальных, образовательных и других особенностей (значимо как в медицине, так и социологии, психологии, педагогике и т. д.), что может быть связано с неумением пользоваться предложенными девайсами или неуверенностью в достаточной защищенности своих данных и т. д. Вопросы разрешения и преодоления личностных рисков зачастую решаются через информирование, инструктирование, грамотное объяснение и желание «идти навстречу» в решении вопроса. Например, если участник биомедицинского исследования принял решение не допускать использования его данных в соответствующих базах, то такие данные должны быть не только скрыты, но и удалены (что в свою очередь требует наличия адекватных алгоритмов и технических решений). Когда исследуемый биологический материал используется без информирования (и разрешения) его владельца, это также можно расценивать как прямое нарушение параметра безопасности и необходимых критериев надежности системы. Классическим примером такой ситуации является история с «бессмертными» клетками HeLa Генриетты Лакс, использованными научным сообществом без ее разрешения (и без информирования семьи), что впоследствии повлекло серьезные проблемы для исследователей. Так, для урегулирования ситуации совместно с представителями семьи Лакс было решено опубликовать расшифровку генома HeLa, ограничив к ней доступ. При желании ознакомиться с этими важными данными ученые должны обратиться в Национальный институт здоровья (США), где их запрос будут рассматривать предметно с точки зрения целей исследования и приоритетов в распространении его результатов с участием в комиссии представителей семьи Лакс («Расшифровка генома «бессмертных» клеток HeLa привела к скандалу в США» – РИА Новости. 07.08.2013 г.).

Безопасность основывается также на устойчивости деятельности систем ИИ, т. е. бесперебойной работе продукта с учетом возможных внешних и внутренних угроз, предполагающей наличие системы самозащиты от таких угроз и влияний. Для соблюдения параметра устойчивости необходимо не только предусмотреть перечень возможных угроз, но также разработать систему соответствующей защиты (например, систему верификации, аутентификации, авторизации пользователей), а также систему оповещения заинтересованных лиц о возникновении незапланированных влияний [12; 13]. В значительной степени устойчивость (и, как следствие, безопасность) определяется и человеческим фактором – действиями пользователей, включая исследователей, технических специалистов, испытуемых и других людей, принимающих участие в работе с системами ИИ. В данном случае для обеспечения устойчивости необходимыми условиями являются тщательное инструктирование, обучение пользователей и постоянный контроль за соблюдением полученных инструкций.

Таким образом, согласие на использование данных, наличие механизмов (алгоритмов), обеспечивающих ограничение доступа к данным, наличие многоуровневой системы защиты пользователей вкупе с защитой от ошибок и искажений ИИ, защита от несанкционированного вмешательства в деятельность системы, возможность установления и снятия ограничений для пользователей, поддержание целостности данных – это обязательные критерии, обеспечивающие безопасность использования ИИ в исследовании. На основании этих критериев должен производиться оценка адекватности использования ИИ как разработчик, так и пользователь, включая исследователя, испытуемого, а также НЭКи, проводящие экспертизу исследований.

Принцип управляемости. Принцип безопасности неразрывно связан с принципом управляемости деятельностью ИИ. Управляемость рассматривается здесь как возможность контроля работы систем ИИ и наличие четко определенной и иерархизированной системы ответственности за действия и результаты, производимые посредством ИИ. Управляемость предполагает следующее:

во-первых, изначальное определение требуемых параметров и установок для применения системы в конкретной сфере. Например, в медицине обязательным условием применения систем ИИ является соответствие таким параметрам, как физическая и психическая безопасность пациента, направленность любых действий на повышение качества его жизни, требование информирования пациента, применение существующих этико-правовых норм, экологичность и др.;

во-вторых, управляемость требует, чтобы системы ИИ (а также их разработчики и пользователи) работали строго в правовом поле, что предполагает принятие на государственном, региональном (и общемировом) уровне специальных рекомендаций и законодательных норм, включающих правовые и этические требования к работе с ИИ;

в-третьих, управляемость системами ИИ обеспечивается созданием и своевременной актуализацией цифрового кодекса (далее – Кодекс), регулирующего правовые отношения участников цифрового пространства. При этом возможность этического самообучения ИИ-систем может быть реализована путем подключения к цифров-версии Кодекса;

в-четвертых, управляемость системами ИИ основывается на формировании системы ответственности. Такая система предполагает иерархию и распределение ответственности между всеми участниками работы с ИИ – от создателя до конкретного пользователя (носителями такой ответственности становятся владельцы баз данных, разработчики, консультанты, тестирующие, поддерживающий работу персонал, а также конечные пользователи). Под конечными пользователями здесь понимаются как собственно организаторы и исполнители исследований, так и участники исследований (испытуемые), которые несут ответственность за точность соблюдения предписанных технологий, соответствие реальных условий использования ИИ обозначенным в предписании (если применимо) и др. Для решения этих вопросов необходимо на национальном уровне продумать систему взаимоотношений вышеперечисленных акторов, четко определить сферу их ответственности, базовые требования и обязанности, предусмотрев анализ ответственности за сбои и погрешности на каждом жизненном цикле ИИ с последующим исправлением ошибок;

в-пятых, управляемость системами ИИ предусматривает механизмы оценки нанесенного системами ИИ вреда (ущерба) с последующей компенсацией. Этот вопрос также должен быть регламентирован правовыми нормами, включающими условия, основания, определение размера и других параметров компенсации [14]. Для биомедицинских исследований проблема компенсаций особенно актуальна. Так, в 1980-х гг. аппарат лучевой терапии Therac-25, разработанный компанией Atomic Energy of Canada Limited «AECL», из-за сбоя в компьютерном кодировании доставлял повреждающие дозы радиации больным раком, что приводило к летальному исходу. Ответственность в этом случае до сих пор обсуждается, поскольку некоторые больницы внедрили собственные обновления систем, которые, возможно, вызвали передозировку. Поскольку в системе ИИ вовлечено много сторон (поставщик данных, разработчик, производитель, программист, пользователь и сама система ИИ), установить ответственность в спорном случае сложно, при этом необходимо учитывать множество факторов [15]. Инцидент с Therac-25 был назван одной из самых серьезных компьютерных ошибок в истории [16], требующих особенного внимания к безопасности ИИ для людей, а также решения проблем оценки и компенсации ущерба.

Принцип объяснимости. Принцип объяснимости связан с наличием принципиальной возможности понимания действий системы, прозрачностью ее алгоритмов, точностью объяснения сути и результатов процесса генерирования результатов [17]. Принцип объяснимости реализуется через критерии прозрачности действий ИИ, основываясь на принципах безопасности и управляемости.

В биомедицинских исследованиях принцип объяснимости чрезвычайно важен для пользователей – как исследователей, так и испытуемых, поскольку паттерн «черного ящика», генери-

рующего результаты, подрывает авторитет исследования и веру в его значимость. Если исследователь не понимает общий алгоритм получения результатов посредством ИИ, он перестает управлять процессом, замечать ошибки, и, соответственно, получаемые результаты становятся исключительно количественными, трудно поддающимися смысловой интерпретации. Ошибки, совершаемые ИИ, достаточно сложны для осознания, поскольку принципиально отличаются от тех, которые мог бы допустить человек. Эта тема, в частности, рассматривается в статье Сяосюань Лю «Привлечение искусственного интеллекта к ответственности» [18], где автор акцентирует внимание на том, что ошибки ИИ человек не всегда может предвидеть, исправить и даже понять.

Парадоксальный пример ошибок в действиях ИИ при диагностике приводят в статье Лорен Окден-Райнер и ее коллеги, проанализировав производительность ИИ для обнаружения перелома шейки бедра [19]. В этом исследовании было выделено несколько «режимов сбоев», т. е. склонность ИИ к периодическим ошибкам в определенных условиях. Среди наиболее существенных «сбоев», связанных с отклонениями от заданных параметров, упоминается случай, когда ИИ «пропустил» сильно смещенный перелом шейки бедра (по мнению авторов, такое изображение даже неспециалист признал бы совершенно не отвечающим норме). Если бы врач не перепроверил полученные данные, была бы допущена серьезная ошибка в диагностике и следующей фазе исследований (лечении), при этом совершенно неясно, кто должен нести ответственность за подобную ошибку.

Не менее катастрофичными могут быть ошибки ИИ при составлении баз данных и статистической обработке результатов. Так, в мае 2018 г. цифровой помощник IBM Watson рекомендовал больным онкологическими заболеваниями неправильные и угрожающие здоровью препараты. Проблема заключалась в применении системой неправильных алгоритмов: вместо обработки данных пациентов и синтеза на этой основе новых идей лечения Watson, как выяснилось, использовал гипотетические данные. Предложения Watson основывались на предпочтениях нескольких врачей, предоставивших данные для разработки системы, а не на реальных выводах, полученных в результате анализа большого количества клинических случаев [20].

Принцип объяснимости работы системы особенно важным становится для распространения результатов исследования в социуме. Отсутствие прозрачности в объяснении подрывает доверие к исследователю (исследовательской команде): с одной стороны, окружая результат ореолом таинственности, с другой – лишая его объективности и научности.

Невозможность объяснить процесс получения результата для структур, осуществляющих этическую экспертизу исследования, как правило, приводит к отрицанию членами этических комитетов ценности самого исследования, отрицанию его научности, требованию перепроверки результатов.

Принцип эффективности. Принцип эффективности предписывает точность и аккуратность данных, собираемых и обрабатываемых системами ИИ. Важным критерием эффективности является соблюдение такого параметра, как «предел знаний», который ограничивает работу систем ИИ теми условиями и целями, для которых система была разработана. Следовательно, самовольное изменение параметров работы систем со стороны пользователей (например, изменение температурного режима, правил эксплуатации, сферы действия и т. д.) угрожает возникновением непредсказуемых ошибок, искажений и, как следствие, несостоятельности данных.

Эффективность работы систем ИИ зависит также и от умения пользователей работать с системой или данными. Иногда некомпетентность приводит к тому, что однозначность результатов анализа ИИ и предиктивные возможности системы чрезмерно преувеличиваются, в результате чего создается «видимость объективности и нейтральности выбора за счет обоснования последнего «непогрешимостью» анализа, проведенного ИИ» [21].

Эффективность работы систем ИИ определяется не только защитой от ошибок и искажений, но также и валидностью анализируемых данных. Так, известный исследователь искусственного интеллекта и его применения в научных исследованиях Сараччи писал, что при построении систем ИИ для нужд современной науки (в частности, для сбора и обработки данных) необходимо ориентироваться не на три «V» (базовые параметры работы с большими данными: где *v* (*volume*) –

размер данных, *v* (*virity*) – разнообразие данных, *v* (*velocity*) – скорость обработки и сбора данных), но также, по мнению Сараччи, современным исследователям необходима еще одна «*v*» (*validity*) – валидность данных [22]. Валидность здесь предполагает максимальную унификацию и стандартизацию подходов и алгоритмов получения данных, постоянный анализ возможных источников искажений. Также валидность требует сбора критичных, важных данных, что определяется исследователями и экспертами в области решения поставленных задач. Хотя вопросы валидности данных прежде всего актуальны для ученых, предлагающих научные и практические выводы на основе анализа данных, однако следует помнить, что при отсутствии валидности сравнение и сопоставление данных как минимум не будут корректными, а конечный продукт не будет отвечать соображениям безопасности и защиты прав пациента.

Вышеприведенные «объективные» этические принципы работы с ИИ (безопасность, объяснимость, эффективность, управляемость) дополняются «субъективными» принципами работы с ИИ, которые, как отмечалось выше, в значительной степени связаны с установками общества и личности, определяемыми уровнем образованности, этико-правовыми нормами, отношением к научным исследованиям, степенью распространенности использования ИИ и т. д.

Принцип справедливости. Справедливость как этический принцип предполагает возможность равного доступа к использованию систем ИИ, равно как и возможность отказа от взаимодействия с ИИ или замены систем ИИ действиями человека. Например, у многих участников исследований может отсутствовать доступ к используемым электронным девайсам/программам, возможно также отсутствие знаний и умений обращения с разными технологиями ИИ при отсутствии помощи (что особенно характерно для людей пожилого возраста) и т. д. Следовательно, для многих участников исследований и пациентов пользование различными медицинскими программами, девайсами и т. д. вплоть до заказа талона на прием к врачу-специалисту на сайте может оказаться затруднительным, а порой и невозможным. Решение вопросов, связанных с равенством доступа, требует от исследователя четко определять уязвимые группы, для которых по разным причинам доступ к использованию систем ИИ может быть осложнен или невозможен, а также изначально предусматривать возможности замены алгоритмов ИИ действиями человека. Например, предоставить возможность заказать талон к врачу-специалисту по телефону (без использования интернета) или, например, вместо применения специального оборудования (девайсов) на дому предусмотреть право осуществлять необходимые действия в поликлинике под наблюдением профильных специалистов и т. д.

Безусловно, с развитием технологий ИИ параллельно развиваются образование общества и возможности людей по использованию ИИ технологий. Тем не менее не следует забывать, что сегодня достаточно большой процент людей не способны, не имеют физической возможности или по своим соображениям не желают работать с системами ИИ, и «вычеркивать» их из исследований (а тем более не оказывать медицинскую помощь) негуманно, нецелесообразно, а порой и критично для адекватности самого исследования. Важно понимать, что реализация равенства доступа испытуемых или пациентов к медицинским исследованиям предполагает наличие специальных знаний и умений у членов исследовательской команды, включающих умение грамотно разъяснять и обучать действиям с необходимыми системами ИИ, информировать об использовании приборов и т. д.

Проблема неравенства доступа к системам ИИ, которая, по сути, становится основой для ИИ-дискриминации, дополняется еще одной неявной проблемой, которую можно определить как «латентная» дискриминация. Последняя связана с когнитивной предвзятостью разработчиков, переносимой на систему ИИ. Если «явная» дискриминация достаточно легко выявляется исследователем или медицинским работником, то работать с «латентной» дискриминацией гораздо сложнее. Для этого исследователю, а также эксперту, консультанту предстоит оценить систему с точки зрения недискриминации, исключить возможность неравенства возможностей и обеспечить работу алгоритмов в соответствии с этическим принципом справедливости. Например, оценить и изменить (если возможно) параметры системы, дающей неравенство доступа, получения результата по признакам расовой, национальной, половой принадлежности, политическим взглядам, религиозным убеждениям, возрасту, социальному и экономическому статусу

или сведениям о частной жизни. По мнению М. Пицци и его коллег, алгоритмы ИИ могут «усиливать существующие неравенства между людьми или их группами, а также усугублять ущемление прав отдельных уязвимых демографических категорий» [21].

Принцип справедливости предъявляет, таким образом, серьезные требования к разработчикам систем ИИ, что предполагает не только анализ системы на наличие или отсутствие дискриминации, но также принятие мер для проверки набора данных, используемых для машинного обучения. Важными задачами становятся создание и применение методик, программных решений, изначально препятствующих возникновению дискриминации, а также адаптация алгоритмов в соответствии с национальными приоритетами и предпочтениями, особенностями ментальности, спецификой национальных (региональных) систем здравоохранения. В докладе о роли ИИ в гуманитарных исследованиях отмечается, что «...система, разработанная в Кремневой долине, но развернутая в развивающейся стране, может оказаться неспособной учесть уникальные политические и культурные особенности данной страны. Разработчик может не знать о том, что в стране N определенные стигматизированные группы населения недостаточно представлены или вовсе остаются невидимыми в наборах данных, и потому не исправит это предубеждение в обучающей модели» [21]. Таким образом, главной этической рекомендацией выполнения принципа справедливости становится внимание разработчиков к специфике модели, с которой они будут работать, включая ее социокультурные особенности. Также необходимы постоянные консультации разработчиков со специалистами той профессиональной сферы, для которой разрабатывается программа (алгоритм). Немаловажным является и требование разработки грамотных и понятных инструкций для работы с конкретной системой ИИ.

Со стороны исследовательской команды для обеспечения принципа справедливости необходимо повышение своей квалификации (уровня знаний) в контексте работы с ИИ, точное следование инструкциям и обязательное предварительное тестирование работы системы с определением рисков в отношении возможной дискриминации участников исследования, с последующим осуществлением соответствующих поправок (если необходимо).

Принцип доверия. Самый сложный в силу своей «субъективности» и относительности – это этический принцип, который основывается на взаимоотношении людей и систем ИИ, – принцип доверия. Этот принцип в отношениях между людьми означает уверенность в человеке, его добропорядочности, искренности намерений и, как следствие, формирование основанных на такой уверенности взаимоотношений. Однако в ситуации с ИИ в системе взаимодействия участвуют как люди, так и «неодушевленный» ИИ, поэтому наряду с «уверенностью» для формирования доверия значимо также субъективное ощущение надежности и принятия.

Понимание доверия предполагает, как минимум, три «среза» взаимоотношений:

доверие потребителей информационных продуктов к разработчикам, предполагающее постоянное и непосредственное решение с ними возникающих проблем;

доверие разработчиков к потребителям, понимаемое как уверенность в ответственности, грамотности, тщательности потребителей в отношении ИИ, строгое следование подготовленным инструкциям;

доверие потребителей к собственно искусственному интеллекту, выражающее чаще всего субъективное эмоциональное отношение человека к ИИ.

Все эти три «среза» доверия базируются на решении нескольких базовых проблем. Прежде всего, это проблема защищенности данных пользователей, которая становится максимально актуальной в контексте биомедицинских исследований, в частности, при использовании парадигмы открытой науки [23]. Фундаментом решения вопроса является четкое определение используемых в современном научном дискурсе понятий «идентифицирующие данные» (данные, через которые можно определить конкретную личность) и «неидентифицирующие», «обратимо анонимизируемые» (обратимо «обезличенные»), псевдоанонимизируемые, персональные и другие данные.

Следующий шаг – формирование возможности использования персональных данных (с разрешения их владельца) и персональной информации ограничено, т. е. построение систем ИИ с учетом ограничения и регулирования доступа к данным. Например, в медицине, в том числе

биомедицинских исследованиях, вопросы доступа к данным пациентов становятся чрезвычайно чувствительными: это связано, в частности, с проблемой возможного доступа к истории болезни, результатам скринингов и другой информации о пациентах со стороны специалистов, непосредственно не участвующих в исследовании или оказании медицинской помощи (сотрудники информационных служб, разработчики, обслуживающий системы ИИ персонал и т. д.), а также врачей-специалистов, медицинского персонала, не имеющих отношения к лечению и обследованию конкретного пациента.

Такая проблема особенно усложняется в связи с использованием больших данных, которые представляют собой огромный массив информации (не только медицинской) о каждом пациенте или испытуемом, аккумулируемой системами ИИ через социальные сети, показатели различных девайсов и даже через телефонные разговоры. Безусловно, для биомедицинских исследований анализ таких комплексных данных абсолютно бесценен, но и риск, что персональные данные станут известны посторонним лицам, достаточно велик. Например, проходя генетическое тестирование, пациент получает очень важную информацию, возможно, определяющую его будущее (наследственные болезни, возможность иметь детей, предрасположенности к заболеваниям и т. д.), но утечка подобной информации в буквальном смысле способна разрушить его жизнь [24]. Таким образом, вопрос регулирования/ограничения доступа к данным дополняется требованием защиты данных через разработку алгоритмов, максимально не подверженных внешним и внутренним рискам.

Второй базовой проблемой, связанной с доверием, является информированное согласие. Информированное согласие (ИС) участника исследований и испытаний – обязательное условие любого биомедицинского исследования. ИС понимается как официальный документ, где на основании предоставленной информации о проводимом исследовании, возможных рисках и потенциальных благах испытуемый дает свое разрешение (согласие) на участие в исследовании [25]. К сожалению, даже тщательно разработанное согласие не всегда отвечает этическим требованиям современной науки и, в частности, требованиям работы с системами ИИ. Например, далеко не всегда испытуемого ставят в известность или спрашивают его волю о дальнейшем хранении и использовании взятого у него биологического материала, данных скринингов и диагностики, что необходимо для создания баз данных. В ИС также должны включаться разъяснения о возможном использовании мобильных приложений в ходе исследований (если применимо), поясняющие, что данные устройства не предназначены для использования в медицинских целях диагностики, мониторинга, лечения, реабилитации, не соответствуют требованиям информационной безопасности в рамках принципов «проектируемой безопасности» и «проектируемой конфиденциальности» [26].

Решение человека, которое он выражает в ИС относительно своих данных и биологических материалов, безусловно, должно неуклонно соблюдаться. Различные виды информированных согласий, в том числе позволяющие назначить «третью сторону» для последующего решения об использовании данных и материалов или ограничить использование, а также уничтожить данные, уже в достаточной степени разработаны в рамках исследовательской этики, поэтому авторы не считают необходимым останавливаться на этом вопросе подробно [27, с. 67–95].

Еще одна проблема, связанная с доверием ИИ, по мнению авторов, заключается в нередко встречающейся «боязни» или неприятия человеком ИИ. Решение этой проблемы лежит в плоскости просвещения и образования общества, создания надежных алгоритмов ИИ и формирования уверенности человека в том, что окончательное решение всегда остается за ним. Речь идет о безусловном праве человека распоряжаться своими данными, принимать решения на основе предоставленной информации и личностных ценностей. В контексте биомедицинских исследований важно акцентировать право и обязанность человека принимать важные решения, касающиеся жизни и здоровья. Примерами таких «человеческих» решений в биомедицине являются те, от которых зависят качество жизни, здоровье и существование в социуме: например, постановка диагноза, решение о паллиативной помощи, распределение ограниченных ресурсов жизнеобеспечения и т. д. ИИ может заменить человеческий труд или аналитическую деятельность, но в большинстве случаев ИИ не может заменить человека «при принятии решений по

особо деликатным вопросам или по проблемам, без устранения которых могут наступить значительные негативные последствия» [22].

Решение проблемы недоверия человека к ИИ, а порой и агрессии в отношении ИИ обеспечивается также через:

прозрачный анализ ошибок, совершаемых ИИ, и их последующую проработку (что тесно связано с реализацией этических принципов безопасности, объяснимости, управляемости);

предусмотренность «человеческих» альтернатив системам ИИ (например, альтернативы использования «ботов»);

оценку профессионализма разработчиков программ в рамках доступа к созданию социально значимого продукта (например, для биомедицинских исследований) вкупе с междисциплинарным обсуждением таких программ;

обязательное внедрение в программы ИИ этических императивов и алгоритмов;

привлечение независимых экспертов для высококвалифицированной оценки систем ИИ.

Выводы. Подводя итоги проведенного исследования, можно сделать вывод о том, что экспликация и анализ основных этических принципов исследований с участием ИИ в контексте безопасного, взаимовыгодного и бесконфликтного взаимодействия ИИ, человека, общества и природы – это первый шаг для осмысления новых вызовов открытой науке, что позволяет создать основу для решения проблем, связанных как с разработкой и обучением ИИ, так и организацией исследований и их этической экспертизой.

Наряду с предложениями, адресованными разработчикам и пользователям ИИ, которые уже были высказаны ранее, считаем необходимым акцентировать внимание на следующих важных положениях:

решение проблемных вопросов, связанных с использованием (совершенствованием) ИИ, должно быть междисциплинарным практически на всех ступенях жизненного цикла ИИ, поскольку междисциплинарность позволяет объединить усилия специалистов в области создания ИИ, специалистов профессиональной сферы, для которой разрабатывается система ИИ, а также специалистов в области этики, права, психологии и т. д., которые обеспечивают адаптацию ИИ для работы в социуме и адаптацию человека к работе с ИИ;

создание этических кодексов для ИИ сегодня необходимо для всех участников взаимодействия «ИИ – человек». Несмотря на то что за последние годы в мире многие организации как локального, так и надгосударственного уровня разрабатывают такие кодексы, проблема остается открытой в силу того, что создаваемые кодексы не основываются на единых принципах, зачастую создаются под ценности и нужды конкретной организации или проекта и не имеют правовой силы [28];

необходимым шагом для развития ИИ и, соответственно, научных исследований с использованием ИИ является создание на основе этических принципов работы с ИИ нормативно-правовой базы, определяющей ответственность, права и обязанности сторон – участников процесса разработки и использования ИИ;

просвещение общества в отношении ИИ, постоянное обучение работе с ИИ участников процесса научных исследований являются необходимым фундаментом развития современной науки, в частности, развития парадигмы «открытая наука»;

важная роль при проведении исследований с использованием ИИ должна отводиться этическим комитетам (или иным этическим структурам). В системе биомедицины – это НЭЖи, обеспечивающие этическую экспертизу исследований, оценивающие соблюдение прав человека и прав природы, соблюдение этико-правовых принципов исследования [9]. К сожалению, существующие в сфере медицины НЭЖи не всегда готовы качественно проводить этическое сопровождение и экспертизу исследований, организованных с применением ИИ, в силу недостаточности знаний об ИИ, неразработанности этических (и правовых) норм исследований с использованием ИИ, отсутствия соответствующих рекомендаций и т. д. В других профессиональных сферах, связанных с исследованиями человека, общества, природы (социология, педагогика, психология, история, экология и т. д.), в нашей стране этические структуры для проведения экспертизы исследований фактически отсутствуют. На сегодняшний день для них не существует ни реко-

мендательных установок, ни правовой базы. Соответственно, реальные практики исследовательских этических комитетов вне медицины крайне редки. Создание и обеспечение работы исследовательских этических комитетов, в частности, подготовка их работы с исследованиями, использующими ИИ, – современное требование развития науки в нашей стране.

Таким образом, необходимо отметить, что тема изучения ИИ по-прежнему остается открытой для изучения, но особенное значение сегодня имеют вопросы этики ИИ. В контексте исследований и испытаний для работы с ИИ современной науке необходимы этические структуры, способные проконсультировать исследователя, провести экспертизу полученных материалов, способствовать распространению научных знаний в обществе и т. д. Разработка парадигмы деятельности таких структур и изучение лучших практик их работы – тема, требующая дальнейшего исследования и развития.

Список использованных источников

1. Предварительный проект Рекомендации ЮНЕСКО об открытой науке [Электронный ресурс] // UNESCO. – Режим доступа: https://unesdoc.unesco.org/ark:/48223/pf0000374837_rus. – Дата доступа: 01.02.2023.
2. Recommendation of the Council on Artificial Intelligence [Electronic resource] // Organisation for Economic Co-operation and Development. – Mode of access: <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>. – Date of access: 07.02.2023.
3. Artificial Intelligence in medicine regulation [Electronic resource] // European Medicines Agency. – Mode of access: <https://www.ema.europa.eu/en/news/artificial-intelligence-medicine-regulation>. – Date of access: 01.02.2023.
4. Statement on artificial intelligence, robotics and «autonomous» systems: Brussels, 9 March 2018 / Europ. Group on Ethics in Science a. New Technologies. – Luxembourg: Publ. Office of the EU, 2018. – 20 p. <https://doi.org/10.2777/531856>
5. Recommendation on the Ethics of Artificial Intelligence [Electronic resource] // UNESCO. – Mode of access: <https://unesdoc.unesco.org/ark:/48223/pf0000380455.locale=en>. – Date of access: 01.02.2023.
6. Ethics and governance of artificial intelligence for health [Electronic resource]: WHO guidance / World Health Organization. – Geneva: World Health Organization, 2021. – Mode of access: <https://www.who.int/publications/i/item/9789240029200>. – Date of access: 07.02.2023.
7. Кодекс этики в сфере ИИ [Электронный ресурс] // Альянс в сфере искусственного интеллекта. – Режим доступа: <https://ethics.a-ai.ru>. – Дата доступа: 01.02.2023.
8. Looking before we leap: expanding ethical review processes for AI and data science research [Electronic resource]: Dec. 2022 // Ada Lovelace Institute. – Mode of access: <https://www.adalovelaceinstitute.org/report/looking-before-we-leap/>. – Date of access: 07.02.2023.
9. Соколичик, В. Н. Роль этических комитетов в обеспечении прав человека при проведении биомедицинских исследований и испытаний в Республике Беларусь / В. Н. Соколичик // Тр. БГТУ. Сер. 6, История, философия. – 2021. – № 1 (245). – С. 146–150.
10. Асеева, И. А. Искусственный интеллект и большие данные: этические проблемы практического использования. (Аналитический обзор) / И. А. Асеева // Соц. и гуманитар. науки. Отечеств. и зарубеж. лит. Сер. 8, Науковедение. – 2022. – № 2. – С. 89–98. <https://doi.org/10.31249/naukoved/2022.02.02>
11. Hagendorff, T. The ethics of AI ethics: an evaluation of guidelines / T. Hagendorff // Minds Machin. – 2020. – Vol. 30, № 1. – P. 99–120. <https://doi.org/10.48550/arXiv.1903.03425>
12. Романова, И. Н. Применение искусственного интеллекта в сфере здравоохранения / И. Н. Романова, О. В. Намов // Устойчивое развитие: исследования, инновации, трансформация: материалы XVIII Междунар. конгр. с элементами науч. шк. для молодых ученых: в 2 т. / отв. ред.: А. В. Семёнов, П. Н. Кравченко. – М., 2022. – Т. 1. – С. 460–468.
13. Vincent-Lancrin, S. Trustworthy artificial intelligence (AI) in education: promises and challenges / S. Vincent-Lancrin, R. van der Vlies. – Paris: OECD Publ., 2020. – 17 p. – (OECD Education Working Paper; no. 218). <https://doi.org/10.1787/a6c90fa9-en>
14. Legal and ethical consideration in artificial intelligence in healthcare: who takes responsibility? / N. Naik [et al.] // Front. Surg. – 2022. – Vol. 9. – Art. 862322. <https://doi.org/10.3389/fsurg.2022.862322>
15. Lynch, J. The worst computer bugs in history: race conditions in Therac-25 [Electronic resource] / J. Lynch // BugSnag: SmartBear. – Mode of access: <https://www.bugsnag.com/blog/bug-day-race-condition-therac-25/>. – Date of access: 07.02.2023.
16. Lynch, J. Therac-25 causes radiation overdoses [Electronic resource] / J. Lynch // BugSnag : SmartBear. – Mode of access: <https://www.bugsnag.com/blog/bug-day-race-condition-therac-25>. – Date of access: 01.02.2023.
17. Reddy, S. Explainability and artificial intelligence in medicine / S. Reddy // Lancet Digit. Health. – 2022. – Vol. 4, № 4. – P. 214–215. [https://doi.org/10.1016/S2589-7500\(22\)00029-2](https://doi.org/10.1016/S2589-7500(22)00029-2)
18. The medical algorithmic audit / X. Liu [et al.] // Lancet Digit. Health. – 2022. – Vol. 4, № 5. – P. 384–397. [https://doi.org/10.1016/S2589-7500\(22\)00003-6](https://doi.org/10.1016/S2589-7500(22)00003-6)
19. Validation and algorithmic audit of a deep learning system for the detection of proximal femoral fractures in patients in the emergency department: a diagnostic accuracy study / L. Oakden-Rayner [et al.] // Lancet Digit. Health. – 2022. – Vol. 4, № 5. – P. 351–358. [https://doi.org/10.1016/S2589-7500\(22\)00004-8](https://doi.org/10.1016/S2589-7500(22)00004-8)

20. Шнуренко, И. Искусственный интеллект на грани нервного срыва / И. Шнуренко // Эксперт. – 2019. – № 1–3 (1103). – С. 39–42.
21. Пицци, М. ИИ в гуманитарной деятельности: права человека и этика / М. Пицци, М. Романофф, Т. Энгельхардт // Междунар. журн. Крас. Креста. – 2021. – № 913. – С. 187–234.
22. Saracci, R. Epidemiology in wonderland: Big Data and precision medicine / R. Saracci // Eur. J. Epidemiol. – 2018. – Vol. 33, № 3. – P. 245–257. <https://doi.org/10.1007/s10654-018-0385-9>
23. Blueprint for an AI Bill of Rights [Electronic resource] // The White House. – Mode of access: <https://www.whitehouse.gov/ostp/ai-bill-of-rights/>. – Date of access: 01.02.2023.
24. Социогуманитарные контуры геномной медицины / отв. ред. Е. Г. Гребенщикова. – М.: ИНИОН РАН, 2021. – 232 с.
25. Голобородько, Н. В. Рекомендации по получению информированного согласия на участие в научном исследовании: учеб.-метод. пособие / Н. В. Голобородько, В. Н. Сокольчик, А. А. Александров. – Минск: БелМАПО, 2020. – 35 с.
26. Добровольное информированное согласие / науч. ред. А. Г. Чучалин, Е. Г. Гребенщикова. – М.: Вече, 2022. – 287 с.
27. Биобанкинг: социально-гуманитарные аспекты / Е. В. Брызгалина [и др.]. – М.: Изд-во МГУ, 2018. – 231 с.
28. Principled artificial intelligence: mapping consensus in ethical and rights-based approaches to principles for AI [Electronic resource]: research publ., 15 Jan. 2020, № 2020-1 / J. Fjeld [et al.] // SSRN. – Mode of access: <http://dx.doi.org/10.2139/ssrn.3518482>. – Date of access: 05.03.2024.

References

1. First draft of the UNESCO Recommendation on Open Science. *UNESCO*. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000374837> (accessed 01.02.2023).
2. Recommendation of the Council on Artificial Intelligence. *Organisation for Economic Co-operation and Development*. Available at: <https://legalinstruments.oecd.org/en/instruments/oecd-legal-0449> (accessed 07.02.2023).
3. Artificial Intelligence in medicine regulation. *European Medicines Agency*. Available at: <https://www.ema.europa.eu/en/news/artificial-intelligence-medicine-regulation> (accessed 01.02.2023).
4. European Group on Ethics in Science and New Technologies. *Statement on artificial intelligence, robotics and «autonomous» systems: Brussels, 9 March 2018*. Luxembourg, Publications Office of the European Union, 2018. 20 p. <https://doi.org/10.2777/531856>
5. Recommendation on the Ethics of Artificial Intelligence. *UNESCO*. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000380455.locale=en> (accessed 01.02.2023).
6. *Ethics and governance of artificial intelligence for health: WHO guidance*. Geneva, World Health Organization, 2021. Available at: <https://www.who.int/publications/i/item/9789240029200> (accessed 07.02.2023).
7. AI Ethics Code. *AI Alliance Russia*. Available at: <https://ethics.a-ai.ru/> (accessed 01.02.2023).
8. Expanding ethical review processes for AI and data science research: December 2022. *Ada Lovelace Institute*. Available at: <https://www.adalovelaceinstitute.org/report/looking-before-we-leap/> (accessed 07.02.2023).
9. Sokolchik V. N. The role of Ethical Committees in securing human rights regarding to biomedical research and tests in the Republic of Belarus. *Trudy BGTU. Seriya 6, Istoriya, filozofiya = Proceedings of BSTU. Issue 6, History, Philosophy*, 2021, no. 1 (245), pp. 146–150 (in Russian).
10. Aseeva I. A. Artificial intelligence and big data: ethical problems of practical use. (Analytical review). *Sotsial'nye i gumanitarnye nauki. Otechestvennaya i zarubezhnaya literatura. Seriya 8, Naukovedenie* [Social and Human Sciences. Domestic and Foreign Literature. Series 8. Science], 2022, no. 2, pp. 89–98 (in Russian). <https://doi.org/10.31249/naukoved/2022.02.02>
11. Hagendorff T. The ethics of AI ethics: an evaluation of guidelines. *Minds and Machines*, 2020, vol. 30, no. 1, pp. 99–120. <https://doi.org/10.48550/arXiv.1903.03425>
12. Romanova I. N., Naumov O. V. Application of artificial intelligence in health. *Ustoichivoe razvitie: issledovaniya, innovatsii, transformatsiya: materialy XVIII Mezhdunarodnogo kongressa s elementami nauchnoi shkoly dlya molodykh uchenykh* [Sustainable development: research, innovation, transformation: materials of the XVIII International Congress with elements of a scientific school for young scientists]. Moscow, 2022, vol. 1, pp. 460–468 (in Russian).
13. Vincent-Lancrin S., Van der Vlies R. *Trustworthy artificial intelligence (AI) in education: promises and challenges. OECD Education Working Paper, no. 218*. Paris, OECD Publishing, 2020. 17 p. <https://doi.org/10.1787/a6c90fa9-en>
14. Naik N., Hameed B. M. Z., Shetty D. K., Swain D., Shah M., Paul R., Aggarwal K. [et al.]. Legal and ethical consideration in artificial intelligence in healthcare: who takes responsibility? *Frontiers in Surgery*, 2022, vol. 9, art. 862322. <https://doi.org/10.3389/fsurg.2022.862322>.
15. Lynch J. The worst computer bugs in history: race conditions in Therac-25. *BugSnag: SmartBear*. Available at: <https://www.bugsnag.com/blog/bug-day-race-condition-therac-25> (accessed 07.02.2023).
16. Lynch J. Therac-25 causes radiation overdose. *BugSnag: SmartBear*. Available at: <https://www.bugsnag.com/blog/bug-day-race-condition-therac-25> (accessed 01.02.2023).
17. Reddy S. Explainability and artificial intelligence in medicine. *The Lancet Digital Health*, 2022, vol. 4, no. 4, pp. 214–215. [https://doi.org/10.1016/S2589-7500\(22\)00029-2](https://doi.org/10.1016/S2589-7500(22)00029-2)
18. Liu X., Glocker B., McCradden M. M., Ghassemi M., Denniston A. K., Oakden-Rayner L. The medical algorithmic audit. *The Lancet Digital Health*, 2022, vol. 4, no. 5, pp. 384–397. [https://doi.org/10.1016/S2589-7500\(22\)00029-2](https://doi.org/10.1016/S2589-7500(22)00029-2)

19. Oakden-Rayner L., Gale W., Bonham T. A., Lungren M. P., Carneiro G., Bradley A. P., Palmer L. J. Validation and algorithmic audit of a deep learning system for the detection of proximal femoral fractures in patients in the emergency department: a diagnostic accuracy study. *The Lancet Digital Health*, 2022, vol. 4, no. 5, pp. 351–358. [https://doi.org/10.1016/S2589-7500\(22\)00004-8](https://doi.org/10.1016/S2589-7500(22)00004-8)
20. Shnurenko I. Artificial Intelligence on the Verge of a Nervous Breakdown. *Ekspert [Expert]*, 2019, no. 1–3 (1103), pp. 39–42 (in Russian).
21. Pizzi M., Romanoff M., Engelhardt T. AI in humanitarian action: human rights and ethics. *International Review of the Red Cross*, 2020, vol. 102, no. 913, pp. 145–180. <https://doi.org/10.1017/S1816383121000011>
22. Saracci R. Epidemiology in wonderland: Big Data and precision medicine. *European Journal Epidemiology*, 2018, vol. 33, no. 3, pp. 245–257. <https://doi.org/10.1007/s10654-018-0385-9>
23. Blueprint for an AI Bill of Rights. *The White House*. Available at: <https://www.whitehouse.gov/ostp/ai-bill-of-rights/> (accessed 01.02.2023).
24. Grebenshchikova E. G. (ed.). *The sociohumanitarian contours of genomic medicine*. Moscow, Institute of Scientific Information in Social Sciences RAS, 2021. 232 p. (in Russian).
25. Goloborod'ko N. V., Sokol'chik V. N., Aleksandrov A. A. (eds.). *Guidelines for obtaining informed consent for scientific research*. Minsk, Belarusian Medical Academy of Postgraduate Education, 2020. 35 p. (in Russian).
26. Chuchalin A. G., Grebenshchikova E. G. *Voluntary informed consent*. Moscow, Veche Publ., 2022. 287 p. (in Russian).
27. Bryzgalina E. V., Alasaniya K. Yu., Varkhotov T. A., Gavrilenko S. M., Ryzhov A. L., Shkomova E. M.. *Biobanking: socio-humanitarian aspects*. Moscow, Publishing house Moscow State University, 2018. 231 p. (in Russian).
28. Fjeld J., Achten N., Hannah H., Nagy A., Srikumar M. Principled artificial intelligence: mapping consensus in ethical and rights-based approaches to principles for AI: research publication, 15 January 2020, no. 2020-1. *SSRN*. Available at: <http://dx.doi.org/10.2139/ssrn.3518482> (accessed 05.03.2023).

Информация об авторах

Сокольчик Валерия Николаевна – кандидат философских наук, доцент. Институт философии, Национальная академия наук Беларуси (ул. Сурганова, 1, корп. 2, 220072, Минск, Республика Беларусь). E-mail: vsokolchik@mail.ru. <https://orcid.org/0000-0002-4975-4052>

Разуванов Алексей Иванович – кандидат медицинских наук, ученый секретарь. Республиканский научно-практический центр медицинской экспертизы и реабилитации (Колодищанский с/с, 93, д. Юхновка, 223027, Минский р-н, Минская обл., Республика Беларусь). E-mail: doc-rai@yandex.by. <https://orcid.org/0000-0001-5033-2933>

Information about the authors

Valerya N. Sokolchik – Ph. D. (Philos.), Associate Professor. Institute of Philosophy of the National Academy of Sciences of Belarus (1 Sarganov Str., Bldg 2, Minsk 220072, Belarus). E-mail: vsokolchik@mail.ru. <https://orcid.org/0000-0002-4975-4052>

Aliaksei I. Razuvanau – Ph. D. (Medical Sciences), Scientific Adviser. Republican Scientific-Practical Centre of Medical Expertise and Rehabilitation (93 Kolodischansky s/s, district of the village of Yukhnovka, 223027 Minsk district, Minsk region, Belarus). E-mail: doc-rai@yandex.by. <https://orcid.org/0000-0001-5033-2933>