

Биоэтика для человечества и искусственного интеллекта

I. Релевантность биоэтики в современном контексте

Появление комплекса технологий, который сегодня принято называть искусственным интеллектом, обострил дискуссии насчёт свойств, присущих исключительно человеку. Научно-технический прогресс традиционно сопровождается распространением мифов и различных иррациональных концепций, и этот феномен всякий раз указывает на недостатки общественного устройства, в котором люди не вполне автономны. В самом деле, существует точка зрения, согласно которой способность философски осмыслить доступную для его восприятия реальность представляет собой ту самую уникальную черту, которая делает человека человеком; антропоцентризм, таким образом — это интерпретация отделённости человека от природы в своей субъектности как его превосходство над таковой. Озабоченность упомянутым вопросом имплицитно признаётсяⁱ авторами Всеобщей декларации о биоэтике и правах человека, принятой в 2005 году, где значится следующее:

“<..> человек обладает уникальной способностью анализировать своё бытие и окружающую его действительность, ощущать несправедливость, избегать опасности, брать на себя ответственность, стремиться к сотрудничеству и давать нравственную оценку происходящему, на основе которой формируются этические принципы <..>”

Автономность, вопреки заявляемому политическому консенсусу, определяется не столько формальной наделённостью человека правами, сколько его умением неординарно мыслить и выбирать, идти или же не идти на компромиссы с другими людьми; смиряться или же не смиряться с укладом общества, членом которого он стал от рождения. Нетрудно убедиться, что самодовольство победителя — говорим ли мы об антропоцентризме или о геополитическом балансе сил — накладывает отпечаток на его образ мышления и поведение, мешая воспринимать положение дел объективно. Именно это обстоятельство во многом привело к ситуации, которую в наши дни справедливо расценивают как критическую: биосфера, будучи подорванной человеческой деятельностью, выходит из равновесия и угрожает людям большими неприятностями, чем обыкновенные погодные катаклизмы.

Биоэтика — направление философской мысли, возникшее и оформившееся как сфера исследований в семидесятых годах прошлого века. В статье в Интернет-энциклопедии Философии, посвящённой биоэтике, говоритсяⁱⁱ, что эта область включает в себя три субдисциплины — медицинскую этику, этику обращения с животными, этику окружающей среды. Строго говоря, если бы экономическая активность последовательно велась в соответствии с моральными принципами, сформулированными ещё древнегреческими философами, или, по крайней мере, осуществлялась бы с

продуманностью, свойственной их манере выражать себя и действовать, мы не были бы сейчас вынуждены в экстренном порядке принимать меры по спасению окружающей среды. Едва ли возможно, придерживаясь античных канонов рассуждения, доказать оправданность намеренного и безнаказанность ненамеренного нарушения природного равновесия. В любом случае, философия, как любая форма познания, подвержена историческим веяниям, и, если её понятийный аппарат в какой-то момент оказывается избыточным, на то есть причины.

В этом эссе я постараюсь сосредоточиться на тех аспектах биоэтики, которые релевантны в контексте автоматизации здравоохранения. Вероятная замена сотрудников этой сферы обученными роботами и использование технологий для диагностики и лечения заболеваний, с одной стороны, требуют подготовки в виде аккумуляции больших массивов отобранных данных и усовершенствования алгоритмов их обработки, с другой — вызывают обеспокоенность как в силу непривычности, так и в связи с высокой сложностью задачи по разработке и внедрению в ИИ правил, аналогичных человеческим морально-этическим нормам. На мой взгляд, необходимость создания такого рода кодекса не самоочевидна, а дополнительным препятствием к его реализации выступает потребность в последующем согласовании пакетов норм для каждого отдельно взятого технологического решения наряду с контролем за соблюдением технических стандартов при их внедрении. Наиболее влиятельные стейкхолдеры — в первую очередь, крупнейшие технологические компании и государства — должны договориться о приоритетах в долгосрочной перспективе, что затруднительно, в частности, потому, что такой ключевой вопрос, как безопасность, сохраняет за собой относительность. Другими словами, стремясь сделать технологию безопасной, мы неизменно задаёмся вопросами: безопасной для кого? Непроницаемой для угроз, исходящих откуда? Что, вообще, считать угрозой?

II. Спорные моменты

Мне импонирует следующая постановка вопроса: какие технологии, наделённые какими характеристиками, в каком темпе и с каким человеческим вмешательством желательно вводить в обращение в сфере медицинских услуг? Допустимо ли собирать личную информацию о здоровье граждан и распоряжаться ей по своему усмотрению без их согласия с целью повышения качества технологий, то есть руководствуясь соображениями общественного блага? Следует ли вручную проверять полученные данные на предмет потенциального закладывания в ИИ механизмов поддержания социального неравенства, например, связанного с более низким качеством медицинских услуг, получаемых людьми не европеоидных рас? Этично ли подвергать пациентов риску, ассоциированному с использованием ИИ, даже с их согласия и при условии доказанной эффективности его применения?

В споре на тему передачи клинических данных сторонним акторам наблюдается деление на два лагеря: одни защищают позицию, что данные находятся в собственности пациента и не могут быть отчуждены, в то время как другие придерживаются той точки

зрения, что данными владеет медицинское учреждение. Доктор Дэвид Ларсон и его исследовательская группа из Школы медицины Стэнфордского Университета выдвигаютⁱⁱⁱ новый подход, основанный на предположении, что, когда речь идёт о вторичном использовании данных, никто не владеет ими в привычном смысле слова. Признавая аргументы в пользу защиты чувствительной информации медицинского характера, учёные выступают за деперсонализацию данных, одновременно подчёркивая, что их реципиенты должны придерживаться высоких этических стандартов. Данный подход подразумевает недопустимость извлечения прибыли из клинической информации как таковой, но допускает финансовые операции с алгоритмами, разработанными на их основе, равно как и возмездную передачу данных для целей исследований. Кроме того, собирателям медицинской информации надлежит взять на себя обязанность не предпринимать попыток к идентификации личностей пациентов. Этичность предложенного подхода подкрепляется следующим тезисом:

“Будучи рассмотрены через предъявленную призму, клинические данные предстают лишь как способ углубиться в фундаментальные аспекты человеческого бытия. Первостепенный интерес представляют не данные сами по себе, а, скорее, физические атрибуты и особенности поведения, заключённые в них.”

Сегодня сложно оценить, насколько реалистично воплощение в жизнь подобной схемы. Прежде всего, слаженное принятие решений в условиях мультитейкхолдеризма осложняется отсутствием терминологического консенсуса; во-вторых, надёжный контроль за их исполнением едва ли возможен в рамках существующих институтов. Известный генетик из Аргентины Виктор Пенчасзаде констатирует^{iv} в материале ЮНЕСКО, подготовленном по случаю двадцатилетия программы Организации по биоэтике, что этот термин распространился на этические вопросы, возникающие в связи с социальным неравенством, не так давно и во многом благодаря деятельности ООН; ранее под биоэтикой подразумевалась, прежде всего, биомедицинская повестка ограниченного числа развитых стран. Академик связывает последствия агрессивной экспансии худших капиталистических практик с отведением правовым и ценностным аспектам биоэтики второстепенного значения; Пенчасзаде утверждает — и сложно с ним в этом не согласиться — что крупнейшие корпорации медицинской индустрии и религиозные организации проталкивают свои интересы и хотят держать идеологический контроль над развитием трендов в биоэтике. Он с сожалением признаёт, что

“Привлечение внимания к вопросам здоровья сегодня, прежде всего, сосредоточено в руках финансистов, озабоченных собственной прибылью <..> и намеренных продолжать классифицировать людей по материальному положению, а вовсе не по требующемуся им лечению, показывая минимальную вовлечённость в действительную ситуацию на всех уровнях.

III. Проблема чёрного ящика и предвзятость автоматизации

Ввиду того, что круг вопросов даже в очерченном мной пространстве взаимоотношений биоэтики и ИИ в здравоохранении чрезмерно велик, дальнейшие

рассуждения будут сфокусированы на двух проблемах, поднятых в статье для АМА Journal of Ethics авторами Майклом и Сьюзан Ли Андерсон^v. Первая — проблема чёрного ящика, неизвестности точных шагов, приведших к принятию ИИ того или иного решения. Вторая — проблема так называемой предвзятости автоматизации, то есть *“склонности людей отдавать предпочтение предложениям автоматизированных систем принятия решений и игнорировать противоречивую информацию, полученную без автоматизации, даже если она верна”^{vi}*. В отношении обоих вопросов исследователи утверждают, что, тогда как при близком к ста процентам показателе распознавания рассматриваемой программой раковых клеток её применение не провоцирует этических противоречий, особенно у пациентов, при допущении ей ошибок разногласия возникают.

Безусловно, проблема чёрного ящика имеет место и должна разрешаться. Самообучающаяся нейросеть не нуждается в человеческом вмешательстве для вынесения диагноза, но сомнения в правдивости сделанных ей выводов обоснованы: тестировщики время от времени фиксируют ошибки в работе нейросети и не в состоянии дать гарантий качества медицинской диагностики с её помощью. На текущем этапе специалисты также не способны рассчитать наверняка, сколько времени потребуется для окончательного устранения неточностей. Есть и другой, представляющийся мне тривиальным, довод в пользу продолжения надзора над эволюцией ИИ-программ — это исследовательский интерес. Собственно, авторы предлагают пойти по пути взаимодополняющего сотрудничества между медицинским специалистом и компьютерной технологией; первый может задаться целью описать те визуальные признаки, по которым система определяет клетки как раковые, постепенно повышая прозрачность алгоритма.

Уместно отметить, что в здравоохранении для всех причастных зачастую присутствуют элементы “чёрного ящика”, и в некотором смысле можно считать рискованной ситуацию, когда пациент без самых базовых представлений о функционировании человеческого организма приходит на консультацию к врачу. Недостаточная внимательность работников медицинских учреждений к особенностям каждого пациента и его потребностям, связанная с образовательными методиками в этой области наравне с устройством институтов и временными ограничениями в предоставлении медицинских услуг, усугубляется невежеством пациентов и невысоким качеством коммуникации между пациентом и врачом. По моему мнению, мотивировать целесообразность внедрения ИИ в сферу здравоохранения тем, что машины менее подвержены физическим, когнитивным и психоэмоциональным колебаниям по сравнению с человеком, неправильно. Не стоит недооценивать важность установления человеческого контакта между врачом и больным, как и в целом принижать значение психологического компонента лечения. Совокупность факторов, которые нужно принять во внимание перед назначением лекарств и восстановительных мер, а также их корреляции, бывают настолько сложны, что возможности для безрисковой автоматизации в этой сфере не так огромны, как представляется.



Я разделяю точку зрения, согласно которой предпочтительно совмещение человеческого интеллекта и машинных мощностей для обеспечения высокого уровня заботы о здоровье людей. При этом не следует забывать о том, что состояние медицинских сотрудников напрямую сказывается на качестве их работы; помимо внесения изменений в образовательные программы и смещения акцента с приобретения знаний на взаимодействие с системами, оснащёнными ИИ, нужно гарантировать медикам хорошие условия труда. Известно, что образ жизни людей сильно влияет на их здоровье, и точно так же, как врач выбирает между назначением более традиционного, лёгкого лечения или, напротив, тяжёлой медикаментозной, инвазивной терапией, мировое сообщество могло бы направить свои усилия на то, чтобы в первую очередь устранить предпосылки к распространению разнообразных заболеваний. Например, культура потребления, трудоголизма, вседозволенности СМИ, недостаточное обеспечение прав детей и другие явления имеют отношение к уровню общественного здоровья. Не исключено, что конструктивное приложение политической воли в перечисленных сферах позволит снизить планку автоматизации, необходимую для предоставления доступа к качественной медицине для как можно большего числа людей. Нагрузки, которым подвергались работники этой сферы во время пандемии, неприемлемы, и одной из задач автоматизации могло бы стать недопущение повторения подобных сценариев в будущем.

В когнитивной психологии описано около двухсот когнитивных искажений; изучение поведения, отклоняющегося от того, что может быть предсказано теорией рационального выбора, привлекает всё больше учёных из разных областей науки. Тем не менее, обучение ИИ, которому предполагается делегировать всё больше полномочий, происходит не только без учёта ошибок, подлежащих исправлению, но и без должного внимания к содержанию датасетов. На этом фоне такое явление, как предвзятость автоматизации, выглядит потенциально опасным; информирование людей о возможностях и ограничениях ИИ должно происходить одновременно с его распространением или даже опережать его. Сегодня люди черпают информацию из источников, которые подбирают сами, а алгоритмы соцсетей предлагают нам материалы, похожие на те, что нам понравились раньше. Таким образом, налицо риск информационной изоляции, когда люди существуют внутри комфортных для них “пузырей” и теряют способность критически воспринимать “чужеродную” информацию, которая вызывает тем больше отторжения, чем более серьёзным подрывом ментального статус-кво угрожает. Наибольшую тревогу вызывает замкнутость круга лиц и компаний, занимающихся разработкой ИИ, ведь именно от их экспертизы зависит, насколько подверженной стереотипам окажется программа на выходе. Я солидарна с американским режиссёром Робин Хаузер в том^{vii}, что, не являясь сторонницей раздутого правительства, разделяю тезис о важности контроля над входными данными для систем ИИ.

IV. Заключение

Биоэтика способна дать ответы на многие вопросы современности, потому что напоминает о принадлежности человека к живой природе; образ жизни людей

отражается на состоянии окружающей среды. Не сдерживая темпы внедрения инновационных технологий, чье функционирование недостаточно транспарентно и чьи бенефициарами является узкая группа лиц, мы развязываем руки представителям циничного капитализма, для которых человеческие тела имеют значение лишь постольку, поскольку задействованы в формировании прибыли. Общественное устройство, в рамках которого подобное считается допустимым, попирает принципы биоэтики в самом широком смысле и навряд ли остановится перед коммерциализацией чувствительной медицинской информации. Силясь предвидеть и устранить негативные побочные эффекты от невнимательного или злонамеренного использования ИИ, мировому сообществу следует уделять повышенное внимание вопросам личной автономии людей, не прикрывая отсутствие прогресса в этой области риторическим фасадом невмешательства в дела суверенных государств. Социально-экономические институты многих стран требуют укрепления, а их слабость не может быть компенсирована внедрением продвинутых технологий в отдельных сферах. Более того, такое положение дел препятствует полноценной реализации принципов биоэтики в мировом масштабе: так, по словам немецкого биолога Вольфганга Хеннига, долгое время проработавшего в Китае, иностранным компаниям не составляет труда обойти существующие в стране ограничения на проведение клинических экспериментов из-за слаборазвитой системы медицинских учреждений в сельской местности^{viii}. Кроме того, крайне важно, чтобы люди, исследующие человеческий мозг на уровнях от молекулярного до клеточного, а также те, чьей профессией является объяснение поведения людей в ситуациях принятия решений, обменивались опытом с разработчиками ИИ. Интеграция этих областей знания поможет предвидеть затруднения, связанные с параллельным участием профессионалов и роботов в выполнении сложных задач.

Список источников

ⁱАкты Генеральной конференции, 33-я сессия, Париж, 3-21 октября 2005 г., т.1:

Резолюции – с. 86. [Электронный ресурс]. URL:

https://unesdoc.unesco.org/ark:/48223/pf0000142825_rus.page=90

ⁱⁱGordon, John-Stewart. Bioethics [Электронный ресурс] : *Internet Encyclopedia of Philosophy*. URL: <https://iep.utm.edu/bioethic/>

ⁱⁱⁱShould clinical data be treated as a public good in the era of AI? *Health Europa*, 24.03.2020

[Электронный ресурс]. URL: <https://www.healtheuropa.eu/should-clinical-data-be-treated-as-a-public-good-in-the-era-of-ai/98842/>

^{iv}Global Bioethics: What For? Twentieth anniversary of UNESCO's Bioethics Programme / под ред. Germán Solinís : UNESCO 2015. [Электронный ресурс].

URL:<https://sustainabledevelopment.un.org/content/documents/1987Global%20Bioethics.pdf>

- ^vAnderson, Michael and Anderson, Susan Leigh. How Should AI Be Developed, Validated, and Implemented in Patient Care? *AMA Journal of Ethics*. 2019;21(2):E125-130. [Электронный ресурс]. URL: <https://journalofethics.ama-assn.org/article/how-should-ai-be-developed-validated-and-implemented-patient-care/2019-02>
- ^{vi}Предвзятость автоматизации – Automation bias [Электронный ресурс] : *Википедия : Свободная энциклопедия*. URL: https://ru.abcdef.wiki/wiki/Automation_bias
- ^{vii}Hauser, Robin. Is Our Future Equitable? *SHRM Executive Network*. Winter 2020. [Электронный ресурс]. URL: <https://www.shrm.org/executive/resources/people-strategy-journal/winter2020/Pages/hauser-feature.aspx>
- ^{viii}Hennig, Wolfgang. Bioethics in China: Although national guidelines are in place, their implementation remains difficult. *EMBO Reports*. 2006;7(9):E850–854. [Электронный ресурс]. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1559670/>