

УДК 101.8

БЕЛИКОВА Е.К.

Московский государственный университет им. М.В. Ломоносова,
г. Москва

СОЦИАЛЬНЫЕ РИСКИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: КУЛЬТУРФИЛОСОФСКОЕ ОСМЫСЛЕНИЕ

Искусственный интеллект (ИИ) вызвал интерес к себе, начиная с самого своего появления. Этот феномен, на первых порах имевший только технологическую природу, в настоящий момент выступает уже как культурное явление. Интерес к нему не утих до настоящего времени и, кажется, только разгорается. Общество с интересом и тревогой следит за всё новыми сферами, которые затрагивает ИИ и в которых он начинает применяться. В связи с широким распространением ИИ появляется множество разнообразных страхов, которые становятся тем значительнее, чем ближе представляется людям появление «сильного» ИИ, обладающего самостоятельным сознанием и полностью повторяющего разум человека. Винить людей в появлении данных опасений не стоит, ведь их «нельзя назвать полностью безосновательными, поскольку в истории уже есть немало примеров того, что технологии негативно воздействуют на жизнь: ухудшают экологию, уничтожают людей в ходе войн, лишают человека приватности, делают его незащищенным перед кибермошенниками и т. п.» [2, с. 19-8]. При этом люди активно используют ИИ: страхи и опасения не становятся достаточно вескими причинами для того, чтобы перестать прибегать к удобствам, предоставляемым различными сервисами, которые функционируют на основе данных технологических разработок. Названное противоречие, по нашему мнению, оказывает значительное воздействие на общественное сознание, имеющее последствия для личностного и общественного постижения результатов применения ИИ, и требует научного анализа.

Цель исследования – попытаться осмыслить представления общества о социальных рисках, которые несёт ИИ, с точки зрения философии культуры. В ходе исследования применялись системно-структурный, ценностный, культурно-исторический методы исследования. Основным подходом стал междисциплинарный, поскольку ИИ в настоящий момент не может быть осмыслен без привлечения методологии различных гуманитарных дисциплин.

Одна из первых угроз обществу, о которой начали говорить в связи с распространением ИИ и вообще компьютеризацией, стала угроза безработицы. Появились опасения, что ИИ способствует автоматизации многих производственных процессов, в связи с чем люди, которые выполняли эти процессы ранее, могут остаться без работы. Кстати, эти опасения пока не оправдались, и внедрение компьютеризированных систем привело, напротив, «не к росту безработицы, а к появлению дефицита высококвалифицированных кадров» [3, с. 246].

Согласно культурфилософскому взгляду на проблему безработицы предлагается обратить внимание на такие факты, как появление у человека благодаря

распространению ИИ свободного времени и освобождение его от неквалифицированного, рутинного, нетворческого труда. Эти процессы, как видим, с точки зрения культуры можно оценить как безусловно положительные.

Ещё один риск, который ИИ вносит в общественное пространство, – угроза кибервойны и кибертерроризма. Этими терминами называют разновидности войны и терроризма, оружием в которых становятся компьютеры и ИИ, происходят хакерские атаки, несанкционированный доступ в чужие компьютерные системы; также в компьютерные системы врага внедряются вредоносные программы и вирусы [6].

Угроза киберпреступности в современном мире также стоит очень остро. Компьютерные правонарушения (похищение информации, средств со счетов и др.) и компьютерное хулиганство (внедрение вирусов в чужие компьютеры) стали частью нашей жизни. Преступники очень изобретательны и постоянно придумывают новые способы нарушения закона, а соответствующие подразделения правоохранительных служб, тоже нередко вооружённые ИИ, им противостоят.

С точки зрения философии культуры, кибервойны, кибертерроризм и киберпреступность представляют собой очередные этапы развития таких негативных социальных явлений, как война, терроризм, преступность, с применением новых, в данном случае технологических инструментов. Для культуры эти процессы связаны в первую очередь не с технологическими изобретениями, а с особенностями человеческой природы, с необходимостью воздействовать на культурные константы, развивать общество и человека так, чтобы количество преступных деяний вообще, в том числе с префиксом кибер-, снижалось. Иначе говоря, в негативных явлениях, связанных с цифровизацией и ИИ, виноват сам человек, а значит, обвинять в этих угрозах ИИ не только бесполезно, но и вредно для общества. В этом случае человек перекладывает на чужие плечи (в данном случае — на ИИ) собственные проблемы, освобождает себя от ответственности вместо того, чтобы попытаться найти пути улучшения ситуации.

Угроза уничтожения человечества системами искусственного разума, который стал «сильным», обрел сознание и решил расправиться со своим создателем-человеком, визуализированы и «проработаны» обществом благодаря художественной литературе, кинопроизводству и телевидению. Среди наиболее известных культурных продуктов, поднимающих данную тему, можно назвать американские франшизы «Терминатор» (1984–2019) и «Матрица» (1999–2021).

С точки зрения философии культуры интересен сам факт появления такого опасения. Человеку кажется, что сильный в физическом и интеллектуальном планах субъект обязательно захочет обрести власть над более слабым. Об этом писал шведский философ Ник Бостром [1, с. 104] и многие другие исследователи, которые в этом случае проецировали собственное поведение на поведение субъекта, который не только является совершенно иным в формальном плане, но даже ещё не создан (появление «сильного» ИИ прогнозируется в течение уже нескольких столетий, но пока так и не состоялось). То, что власть обязательно

связана с желанием деструктивно воздействовать на тех, кто находится в подчинении, признаётся в подобных работах сущностным свойством в первую очередь человека, а не ИИ.

Менее разрушительной для человечества, но не менее актуальной для его существования является «угроза приватности» [5, с. 70]. Для того чтобы системы ИИ функционировали более эффективно, человек вынужден отказаться от части своей приватности, от закрытости своей личной информации. Передвижения человека, его действия, мысли и чувства сегодня невозможно держать в тайне, они могут быть доступны недоброжелателям. Под угрозой личная информация, интимные сведения, семейные секреты и многое другое. Эта информация легко может оказаться в руках недоброжелателей, причём представляющих как себя лично, так и государственные, властные структуры, стремящиеся к контролю над гражданами.

С точки зрения философии культуры за данными возможностями и действиями скрываются опять же человеческие негативные качества: либо любопытство к чужой жизни, неспособность к уважению личного пространства человека, либо желание использовать личную информацию для унижения подвластных тебе людей, для контроля и подчинения. Здесь мы снова видим, как на ИИ проецируются отрицательные намерения человека: на деле искусственный интеллект лишь используют как инструмент.

Справедливость требует, чтобы мы назвали не только риски, которые несёт обществу ИИ, но и положительные аспекты его общественного функционирования. Именно благодаря системам, действующим на основе ИИ, значительное развитие получили такие ценности, как милосердие и взаимопомощь: люди с готовностью пользуются различными сервисами, с помощью которых можно легко и посредством внесения небольших сумм заниматься благотворительностью, а также становиться волонтерами и бескорыстно помогать людям. Активно развиваются благотворительные организации, технологически функционирующие на основе ИИ-систем. Сегодня большее количество людей проявляет милосердие, и это связано с развитием традиционной ценности – коллективизма.

В современных социальных сетях и мессенджерах одной из наиболее востребованных является функция создания группы, в которой можно организовать коллективное общение: семейное, производственное, по интересам и др. В организации таких групп принимает участие ИИ, отслеживающий интересы пользователей и предлагающий им варианты сообществ и объекты коммуникации. В результате человек получает новые возможности для более интенсивного развития такая ценность, как уже упомянутый коллективизм, а также семейные ценности, поскольку сообщества дают возможность укреплять семейные связи.

Наиболее инновационной и даже не до конца освоенной человечеством ценностью выступает сегодня виртуальная реальность. Это синтезированный с помощью компьютера и ИИ особый мир, существование в котором схоже с существованием в реальном мире. Полноценная виртуальная реальность (полное погружение в виртуальное пространство) пока только создаётся, однако само пространство интернета, доступное большинству жителей земного шара, уже можно назвать виртуальным пространством. Людей привлекает в нём возможность

мгновенного перемещения в любую точку земного шара, доступность разнообразной информации, возможность найти как интересную работу, так и развлечения. Мы оцениваем доступность для человека этой новой реальности положительно потому, что считаем необходимым расширение человеком пространства своего пребывания. Чем больше возможностей открывается перед человеком, тем более активно он развивается.

И угрозы, которые человек предполагает со стороны ИИ, и положительные следствия использования ИИ исходят, по нашему мнению, от самого человека, а не от каких-то особенных свойств, которыми ИИ обладает, а человек – нет. В противоположности положительного и отрицательного, исходящего от ИИ, проявляются противоречия собственной человеческой натуры — то есть и лучшие, и худшие качества, которые есть у человека и которые влияют как на культуру вообще, так и на её разновидность, составляющую, слой — культуру искусственного интеллекта. Распространение ИИ позволяет нам лучше видеть такие антиномии человеческой сущности, как стремления: быть добрыми и общаться – властвовать и довлекать; бескорыстно помогать людям – подчинять людей своей воле; строить гуманное общество и систему отношений в нём – нарушать закон и совершать преступления против общества и его членов и т. д. ИИ не создал ни плохого, ни хорошего в данных диадах; он только воплощает представления людей о плохом и хорошем. К примеру, ИИ может как быть источником террористической угрозы, так и использоваться для того, чтобы такие угрозы выявить, устранить или предотвратить.

Вопрос о том, что превалирует, социальные угрозы или социальные преимущества, исходящие от ИИ, никогда не будет решён однозначно. Возможны технооптимистическая и технопессимистическая точки зрения. Ряд исследователей-технопессимистов с уверенностью заявляет: «Негативные аспекты применения технологий искусственного интеллекта способны превысить общий положительный эффект технологий искусственного интеллекта» [4, с. 15]. Для нашего исследования важно то, как соотношение рисков и преимуществ может быть оценено в культурфилософском плане.

Культура динамична и не может остановиться в своём развитии. Стабильные компоненты культуры постоянно пополняют её новые элементы, возникающие в результате развития различных сфер человеческой жизни, в том числе технологической сферы. При этом культура достаточно константна и имеет устойчивое ядро, пошатнуть постоянство которого не просто очень сложно, а невозможно. Наблюдаемые изменения касаются в первую очередь периферийных сфер культуры и исключительно медленно проявляются в ядерной ее части.

Итак, в современном социальном пространстве за время, прошедшее с появления и распространения ИИ, оформился комплекс страхов, отражающих опасность ИИ для человечества. Поскольку эти риски предполагают наличие опасности для всего общества и его отдельных членов, их принято называть социальными (в противовес культурным, которые опасны более не жизни и существованию людей, а их культуре и ценностям). К основным социальным рискам ИИ можно отнести угрозу безработицы, угрозы кибервойны, кибертерроризма и

киберпреступности, угрозу уничтожения человечества системами искусственного разума, угрозу приватности. С культурфилософской точки зрения все данные угрозы исходят не от ИИ, а от человека, отражая те негативные стороны его природы и оформившихся особенностей его существования в обществе, которые сложились культурно и исторически. В этой связи необходимо помнить о тех преимуществах, которые несёт ИИ и которые также отражают человеческие черты: облегчение процессов милосердия и взаимопомощи, развитие коммуникативных ресурсов, помощь в создании группы и развитии индивидуальности человека, укрепление семейных связей, а также развитие виртуальной реальности, которая повышает информационную компетенцию человека.

ИИ является продуктом человеческой цивилизации, а культура ИИ – частью человеческой культуры. Поэтому все опасения, которые высказывает человек по поводу устройств с ИИ, можно смело называть опасениями в свой собственный адрес. Люди, таким образом, боятся не ИИ, а самих себя, тех саморазрушительных тенденций, которые они априори предполагают в структуре собственной личности.

Список литературы:

1. Бостром Н. Искусственный интеллект. Этапы. Угрозы. Стратегии / пер. с англ. С. Филина. М.: Манн, Иванов и Фербер, – 2016. – 404 с.
2. Крылова М.Н. Социальные угрозы, которые вызывает искусственный интеллект, и реакция на них общества // *Studia Humanitatis*. – 2024. – № 1. – С. 19.
3. Крюков А.Н. Влияние искусственного интеллекта на занятость и безработицу в Российской Федерации // XXXVI Международные Плехановские чтения: сб. ст. участников конф. – М.: РЭУ им. Г.В. Плеханова, 2023. – С. 243-247.
4. Романова И.Н. Внедрение технологий искусственного интеллекта: анализ вероятных рисков и возможных угроз // *Материалы Ивановских чтений*. – 2021. – № 4 (35). – С. 15-18.
5. Самойлов А.В. Технологии искусственного интеллекта: возможности или угрозы // *Журнал высоких гуманитарных технологий*. – 2023. – № 3 (3). – С. 67-71.
6. Clauberg R. Cyber-physical systems and artificial intelligence: chances and threats to modern economies // *World Civilizations*. – 2020. – Т. 5. – № 3-4. – P. 107-115.