

Практики цифровизации

УДК 101.1:316

Е. В. Сажиевко

Эволюционные перспективы развития искусственного интеллекта

Аннотация:

В статье рассматриваются различные подходы к пониманию будущего взаимодействия человека и технологий. Проводится оценка эволюционных перспектив развития искусственного интеллекта, в том числе, высказываются предположения о том, какие перспективы более реалистичны. Анализируются проблемы, которые может породить его развитие.

Ключевые слова: искусственный интеллект, человек, технологии

Об авторе: Сажиевко Екатерина Владимировна, кандидат философских наук, Государственный университет «Дубна», ассистент кафедры социологии и гуманитарных наук факультета социальных и гуманитарных наук; эл. почта: cuore-87@mail.ru

Активное внедрение технологий в различные сферы жизни все больше погружает человека в созданную им самим среду, частью которой становится искусственный интеллект. Эта среда все больше отдаляет человека от биологического естества.

Сегодня под термином «искусственный интеллект» понимается «результат деятельности человека, который представляет собой сложную совокупность коммуникационных и технологических взаимосвязей, обладающую способностью логически мыслить, управлять своими действиями и корректировать свои решения в случае изменения внешних условий» [10, с.41].

Это технологии, совмещающие способность к логическому мышлению, решению творческих задач, пониманию устной и письменной речи, ведению осмысленного диалога,

обучению, самообучению и возможности принимать решения на основе постоянно обновляющихся знаний о ситуации и мире в целом. Конечно, сегодня существуют технологии, способные написать музыкальную композицию, обыграть гроссмейстеров в шахматной партии, понимать и воспроизводить устную речь и т.д. Есть и более сложные технологии, способные выполнять несколько задач. Несколько лет назад всеобщее внимание было приковано к человекоподобному роботу Софии с искусственным интеллектом, способным к обучению, но контролируемым разработчиками. Он участвовал в крупнейших мировых форумах и давал интервью журналистам. Осенью 2017 г. робот получил подданство Саудовской Аравии. Подданство и гражданство раньше было исключительным правом человека, но сегодня это уже не совсем так. Сам механизм идентифицирует себя как женщину, но осознает, что он/она именно робот, а не человек. Он может шутить, запоминать информацию, воспроизводить и использовать ее в последующем. Все же пока нет программ и машин, позволяющих совмещать все эти возможности, имманентно присущие человеку. Главное, что до сих пор не созданы технологии, способные к автономному самообучению без вмешательства создателя.

В современном мире нет единства в отношении быстрого технологического роста в целом и развития искусственного интеллекта в частности. Можно выделить три основных подхода к будущему взаимодействию человека и технологий [2]:

1. Слияние технологий и биологического вида *homo sapiens*.
2. Технологии продолжают делать нашу жизнь более простой, но человек остается человеком.
3. На смену человеку полностью придут машины, которые либо подчинят его, либо полностью вытеснят.

Особую опасность для человека могут представлять технологии искусственного интеллекта, основанного на принципе самообучения. Если не вводить в алгоритмы функционирования специальных ограничений, то саморазвитие искусственного интеллекта может привести к тому, что он довольно быстро превзойдет интеллект человека, который, конечно, тоже трудно назвать полностью естественным. Тем не менее есть вероятность, что именно эти технологии можно использовать для «когнитивного улучшения» человека. Современная медицина активно использует новейшие разработки для повышения качества жизни людей, страдающих тяжелыми заболеваниями. Возможно, искусственный интеллект

или его отдельные элементы со временем удастся использовать для улучшения жизни людей с ограничениями в когнитивной сфере.

И все же многие ученые, изобретатели и общественные деятели с опаской говорят о перспективах развития искусственного интеллекта. Например, Илон Маск, апологет научных исследований в разных областях, особенно в сфере развития и освоения космоса, полагает, что он представляет собой экзистенциальную угрозу для всего человечества. Самообучающаяся система сможет быстро превзойти человека, если процесс самообучения не будет ничем сдерживаться. Его мнение разделял, например, Стивен Хокинг. И Маск, и Хокинг – люди, которых крайне трудно обвинить в технофобии. Кроме того, часто отмечается, что «даже если система, использующая искусственный интеллект, управляется человеком, последний может иметь ограниченный контроль над всеми возможными действиями и реакциями системы, равно как и разработчики не всегда могут предугадать поведение самообучающихся систем» [8, с. 203].

Один из способов обеспечения безопасности может быть связан «со стремлением к моделированию безопасного поведения и мышления человека (при использовании нейрофизиологии и психологии), т.е. физиологических основ работы мозга, которые трансформируются в цифровую модель или искусственный интеллект» [7, с. 64]. Но есть ли техническая возможность перенести эти аспекты функционирования человеческого мозга на электронный носитель?

В связи с намечающейся угрозой со стороны развивающейся технологии также остро встает вопрос ответственности и правового регулирования отношений, возникающих в связи с ее использованием. Наибольшее внимания заслуживает то, кто будет нести ответственность за действия систем, использующих искусственный интеллект: его создатель, человек, занимающийся его эксплуатацией или сама технология. Уже сегодня в отечественной юридической литературе можно встретить понятие «электронное лицо» [10].

Подобные изменения представляют собой экстраполяцию роста технологической мощи как одного из векторов или мегатрендов эволюции. А.П. Назаретян писал, что «на протяжении этого столетия планетарной цивилизации предстоит либо очередной, причем беспримерный по крутизне виток «удаления от естества» (что во всех переломных эпохах обеспечивало преодоление антропогенного кризиса), либо столь же беспримерный по масштабу обвал» [6, с. 191]. При таком подходе представляется, что не придется долго

ждать целостных и глубоких изменений, спровоцированных технологическим развитием. Все большее внедрение технологий происходит на наших глазах ускоренными темпами, что во многом вызвано пандемией коронавируса Covid-19. Например, система образования уже претерпела значительные изменения во многих странах мира, хотя предпосылки такого изменения складывались на протяжении предыдущих десятилетий.

Технологии, которые сегодня столь активно разрабатываются, уже в ближайшем будущем могут привести к изменению общества в целом, могут исчезнуть старые профессии, а на их смену придут новые. Так происходило во все эпохи, связанные с технологическими прорывами. Уже не раз в истории человечества озвучивались тревожные прогнозы о массовой безработице. Ситуацию, складывающуюся сегодня, нельзя считать исключением: «в ближайшие годы, накапливавшиеся десятилетиями изменения, характерные для технологий, нацеленных на трудосбережение и замену человека машинами и автоматами, приведут к радикальному изменению производства и рынка труда» [1, с. 173]. Конечно, все это случится. Но рост безработицы, предсказываемый автором, вероятнее всего, будет временным. Рынку труда понадобится время, чтобы подстроиться под изменившиеся условия, например, под автоматизацию многих обязанностей, сегодня выполняемых человеком. Поскольку часть полномочий постепенно передается технологиям есть вероятность снижения значимости даже тех профессий, которые требуют высокого уровня образования, например, школьного учителя или врача (хотя сейчас медицинские работники чрезвычайно востребованы во всем мире). Процесс «технологизации» даже творческих профессий (современные программы могут писать музыку, нейросети помогают создать продолжения популярных циклов книг, хоть эти продолжения пока далеки от совершенства), скорее всего, будет продолжаться и впредь. Но смогут ли информационные технологии и искусственный интеллект заменить человека в полном объеме?

Из-за стремительного развития технологий выглядят реальными как сценарии обвала цивилизации и начала нисходящей ветви эволюции, так и прогрессивные сценарии развития, связанные с переходом системы на новый уровень сложности (совершенствование робототехники, развитие искусственного интеллекта, космическая экспансия, даже создание интерфейса «мозг-компьютер»). Например, Р. Курцвейл предполагает, что появятся технологии, способные напрямую подключать мозг человека к облаку, в котором хранится весь массив информации [4].

Нельзя исключить и варианты развития, при которых человек останется биологическим существом, который тем не менее может значительно измениться в результате улучшения генома, замены внутренних органов или биопротезирования. Пожалуй, из всех подходов к будущему это один из самых реалистичных и сохраняющих «надежду» на сохранение человека. Даже искусственный интеллект, опасность бесконтрольного развития которого сегодня вызывает множество дискуссий, действительно, представляет собой вполне реальную проблему, несмотря на то что его еще не существует, по крайней мере, в смысле, вкладываемом в это понятие в современном мире.

Человечество как биологический вид постепенно вырождается. Естественный отбор почти полностью заблокирован гуманистической культурой, из-за этого последующие поколения постепенно становятся генетически слабее предыдущих [3]. Дальнейшее выживание нашего вида требует применения все новых технологий в медицинских целях, что, в свою очередь, может привести к продлению жизни и повышению ее качества у людей с серьезными заболеваниями. Хотя многие исследования в этой сфере запрещаются по этическим соображениям, потому что даже в тех научных отраслях «где до сих пор не завершилась «демаркация границ» морально допустимого, наделение человека свойствами, совершенно чуждыми его природе, по-прежнему остается табу» [5, с. 66]. Только нормы этики, которыми руководствуются в научной среде, остаются важным ограничителем исследований в некоторых сферах (например, запрещены эксперименты по клонированию человека). Близкие по духу советы по этике должны, на наш взгляд, регулировать исследования и разработки в сфере искусственного интеллекта. Этика и открытость этих разработок были включены в список принципов полезного искусственного интеллекта в ходе Асиломарской конференции в 2017 г.

Существуют и другие перспективы для человечества. Довольно часто звучат слова о том, что необходимо отказаться от телесности как основы идентичности человека, симбиоза человека и машины, все-таки более «правдоподобным представляется сценарий сохраняющего симбиоза, без которого не может обойтись биологически слабеющее человечество» [6, с. 193].

Новейшие технологии способны улучшить качество жизни даже тяжелобольных людей. Тем не менее, вероятность полной замены человека аватаром довольно низка. При

этом остается неясной цель таких устремлений. Эмоциональная сфера, ее специфика и развитие – одна из фундаментальных черт индивидуальности. Стоит ли возможное бессмертие рациональной части личности ее полной утраты или серьезного искажения?

И все же идеологи трансгуманизма полагают, что в обозримом будущем появятся искусственные механизмы, способные выйти «на передний край эволюции разума и поведут за собой (а может быть полностью поглотят) все более отстающих человекoв» [9, с. 601]. Это уже не простое создание роботов, это попытки определенной части изобретателей создать сверхчеловека, способного прийти на замену homo sapiens. Интеллект этого нового вида может превзойти человеческий, но достаточно трудно всерьез рассуждать о наличии личности у аватаров. И все же такая идеология не оказывает серьезного воздействия на наших современников, хотя и зародилась больше столетия назад.

Вероятно, ближайшие технологические прорывы будут связаны с развитием медицины и робототехники. Интерес к медицине и исследования в этой области сегодня будут, пожалуй, самыми актуальными, так как пандемия в очередной раз наглядно продемонстрировала хрупкость человеческого организма и его незащищенность перед естественной средой.

Проблема искусственного интеллекта довольно часто обсуждается в контексте проблемы сингулярности. Например, в 2012 г. была опубликована антология «Singularity Hypothesis: A Scientific and Philosophical Assessment», в самой структуре которой выделено два ключевых подхода к сингулярности. В рамках первого подхода сингулярность напрямую связывается с развитием искусственного интеллекта. Второй основан на идеях трансгуманизма и связан с созданием постчеловеческого интеллекта. Апологеты искусственного интеллекта акцентируют внимание на его преимуществах: увеличении вычислительных возможностей, скорости коммуникации, редактируемости, регулировании целей, точной воспроизводимости при каждом повторении, улучшенной рациональности [12]. Их оппоненты считают, что роботы не обладают очень важной способностью – многозадачностью, они могут концентрироваться на решении только одной задачи. По этой причине возможности естественного интеллекта априори превышают возможности искусственного [11].

В результате остается три варианта действий:

Во-первых, прекращение исследований в этой области и запрет на совершенствование подобных технологий, аналогично с запретом на клонирование человека. Но здесь может возникнуть множество трудностей, поскольку за такими разработками трудно уследить. Для продолжения исследования достаточно группы энтузиастов. Кроме того, тысячелетние тренды невозможно изменить простым управленческим решением: «все попытки задержать развитие технологий из тех или иных социальных соображений (в основном для предотвращения безработицы) заканчивались провалом» [1, с. 181]. Ни одно современное государство или международная организация не могут этого совершить.

Во-вторых, разработка элементов искусственного интеллекта, которые позволят улучшить качество жизни человека, произведут революцию, например, на рынке труда, облегчат бытовые условия и т.д., при контроле, в том числе, этическом над всеми исследованиями в этой области. Контроль со стороны научного сообщества и институтов власти, вероятно, предотвратит развитие событий по неблагоприятному сценарию, о возможности которого предупреждают многие ученые и изобретатели.

И, наконец, разработка полноценного искусственного интеллекта в полном объеме, т.е. создание технологий, способных к проведению логических операций и решению задач творческого характера, а также к совершению любых других действий, связанных с интеллектуальной деятельностью.

От выбора одного из этих вариантов будут зависеть и эволюционные перспективы искусственного интеллекта. Все же он представляет собой не прямую угрозу. Человек может утратить место «двигателя» эволюции, которое может быть занято искусственным интеллектом как новым эволюционным лидером. В то же время все разработки в этой сфере можно использовать и для развития самого человека, если проявить осторожность и не поддаваться ни панике, ни эйфории.

Библиографический список:

1. Акимов А.В. Робототехника и трудосберегающие технологии: перспективы воздействия на социально-экономическое развитие // Историческая психология и социология истории. 2017. №1. С. 173-192.
2. Баррат Дж. Последнее изобретение человечества: искусственный интеллект и конец эры Homo Sapiens. М. : Альпина нон-фикшн, 2019. 396 с.

3. Бочков Н. П. Генетика человека. Наследственность и патология. М. : Медицина, 1978. 380 с.
4. Курцвейл Р. Эволюция разума. Как расширение возможностей нашего разума позволит решить многие мировые проблемы. М. : Издательство «Э», 2015. 352 с.
5. Майло П. Что день грядущий нам готовил? М. : ЛомоносовЪ, 2011. 312 с.
6. Назаретян А.П. Цивилизационные кризисы в контексте Универсальной истории. М. : PerSe, 2001. 240 с.
7. Соколова С.Н. Искусственный интеллект и безопасность общества // Вестник Полесского государственного университета. Серия общественных и гуманитарных наук. 2016. №1. С. 63-68.
8. Шестак В.А., Волеводз А.Г. Современные потребности правового обеспечения искусственного интеллекта: взгляд из России // Всероссийский криминологический журнал. 2019. №2. С.197-206.
9. Эпштейн М. Знак пробела: О будущем гуманитарных наук. М. : НЛО, 2004. 864 с.
10. Ястребов О.А. Правосубъектность электронного лица // Труды института государства и права РАН. 2018. Т.13, № 2. С. 36-52.
11. Bray D. Brain versus Machine // Singularity hypotheses. A Scientific and Philosophical Assessment. Berlin Heidelberg: Springer-Verlag, 2012. P. 269-279.
12. Muehlhauser L., Salamon A. Intelligence Explosion: Evidence and Import // Singularity hypotheses. A scientific and philosophical assessment. Berlin Heidelberg: Springer-Verlag, 2012. P. 15-40.

Sazhienko E. Evolutionary prospects for the development of artificial intelligence

The article examines some approaches to the understanding of the future interaction between human and technology. The evolutionary prospects for the development of artificial intelligence are assessed, including assumptions about realistic and unrealistic prospects. In addition, the article analyses how the artificial intelligence can give rise to the new problems.

Keywords: artificial intelligence, human, technology