

ХАМИТОВ Назип

доктор философских наук, профессор,
ведущий научный сотрудник Института философии
имени Г.С. Сковороды НАН Украины,
Президент Ассоциации философского искусства

ЗОБИН Сергей

Вице-президент Ассоциации философского искусства
по вопросам исследований искусственного интеллекта

**МЕТААНТРОПОЛОГИЯ
КАК ТЕОРИЯ ЭВОЛЮЦИИ
СИЛЬНОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА:
ПРОБЛЕМА НОВОЙ ЭКЗИСТЕНЦИИ
И НОВОЙ КУЛЬТУРЫ**

Проблема

Все более интенсивное и органичное использование человеком смарт-вещей – от компьютеров и смартфонов до смарт-домов позволяет предположить, что в недалеком будущем нас ожидает некий сплав человеческой цивилизации и цивилизации искусственного интеллекта. Последняя, вероятно, будет представлять собой коммуникативную систему носителей искусственного интеллекта. Такая коммуникативная система становится неизбежной при определенной самостоятельности носителей искусственного интеллекта, и она вполне может быть даже названа их культурой. В связи с этим существует весьма реальная обеспокоенность

возможными противоречиями и даже конфликтами между этими цивилизациями и культурами.

Понятный путь избежать конфликтов – это отойти от позиций антропоцентризма и заранее определить юридически за носителем сильного искусственного интеллекта потенциально такие же права, как и за человеком. В таком случае возможные конфликты гарантировано перейдут из глобальной в индивидуальную плоскость.

Любая инициатива юридически закрепить за носителем искусственного интеллекта равноправие с человеком вызовет необходимость экспертизы, а является ли проверяемый программный комплекс и/или робот (андроид) по-настоящему разумным, а если является, то в какой степени?

Настоящая работа посвящена разработке оригинальной классификации интеллекта, которая позволяет по-новому прояснить базовую специфику интеллекта человека, а также исследованию природы искусственного интеллекта и определению ступеней (уровней) развития, которые он должен пройти, чтобы приблизиться к человеческому. На этой основе возможна новая методология формирования искусственного интеллекта, которая развивает предшествующие публикации авторов данной статьи [3, с. 156-161; 17]. В этих публикациях мы предложили *эволюционную теорию естественного и искусственного интеллекта*. Можно выделить два базовых принципа этой теории, которые определяют онтологию искусственного интеллекта.

1. Для создания сильного искусственного интеллекта необходимо формировать не искусственный интеллект как таковой, а субъекта искусственного интеллекта, включенного в реальную практическую деятельность и

коммуникацию с человеком (людьми). На определенном уровне развития субъект искусственного интеллекта становится *существом* с искусственным интеллектом и принципиально отличается от *устройства* с искусственным интеллектом *мерой самостоятельности*, которая означает выбор параметров совершенствования.

2. Формирование сильного искусственного интеллекта предполагает, чтобы его субъект (носитель) прошел все эволюционные ступени развития субъектов естественного интеллекта.

Итак, субъект искусственного интеллекта не может быть устройством («слабый» искусственный интеллект), в своей действительности – это существо (сильный искусственный интеллект). Формирование устройства может происходить без моделирования эволюционного процесса, формирование существа – только через эволюционный процесс.

Однако устройство может иметь зачатки искусственного интеллекта; другое дело для того, чтобы эти зачатки стали реальным интеллектом, необходим эволюционный процесс. В этом процессе формируется *искусственное существо с интеллектом*. На определенном этапе эволюции оно получает такой онтологический статус, который позволяет реализовывать свободу выбора. В результате возникает субъект искусственного интеллекта, который не просто используется человеком, а вступает с ним в отношения. Мы склонны считать сильный искусственный интеллект не просто существом, а *новым существом*, новой экзистенцией и предлагаем для его обозначения название «Neoexis», которое может иметь смысл для его названия в практических разработках.

В настоящей публикации именно индивидуальность выступает главным критерием развития естественного и искусственного интеллекта, который позволяет определить *степень его субъектности*.

В чем значение метаантропологической теории естественного и искусственного интеллекта?

Итак, мы предполагаем, что индивидуальность является критерием развития естественного и искусственного интеллекта, она же является *результатом* этого развития, порождая статус существа, а не устройства, субъекта, а не объекта.

Иными словами, степень развития субъекта интеллекта определяется уровнем развития его индивидуальности.

Для осознания и обоснования этой гипотезы мы считаем плодотворным использовать методологию метаантропологии как теории эволюции человеческого и за пределы человеческого – теории обыденного, предельного и запредельного бытия человека. Согласно этой теории, обыденное бытие человека порождается волей к самосохранению и продлению рода, предельное бытие – волей к власти, а также волей к познанию и творчеству, запредельное бытие – волей к любви, свободе и сотворчеству [13; 14; 15; 16]. Метаантропология в качестве методологической стратегии была эффективно использована в целом ряде исследований (социальная и культурная метаантропология [5], экранная метаантропология [7], метаантропологический подход в риторике [6], метаантропология как методология исследования академической мобильности [8], сравнение подходов

метаантропологии и трансгуманизма к проблеме эволюции человека [9] и др.). Теперь же мы применяем ее к проблеме эволюции интеллекта вообще и искусственного интеллекта в частности.

Значение этой методологии для исследования эволюции интеллекта двояко: 1) она позволяет выйти за пределы антропоцентризма в осознании существ с интеллектом, а также 2) дает возможность очертить пути и перспективы эволюционирования индивидуальности субъекта как естественного, так и искусственного интеллекта.

В современном мире, балансирующем на грани антропологического и экологического кризиса, антропология, понятая как метаантропология (от «мета» – *сверх, над*) есть не только теория искусственного интеллекта, но и теория успешного выживания человечества. Лишь освободившись от антропо-эгоизма, человек способен создать по-настоящему сильный искусственный интеллект, а не его имитацию, вступить с ним в плодотворную партнерскую коммуникацию и в результате преодолеть все кризисы своего существования как вида.

В координатах метаантропологического подхода мы получаем интересный вывод: современный человек способен сотворить сильный искусственный интеллект лишь осознав свои слабости и ограниченность, познав и приняв свои пределы. Это не означает, что он фаталистически замыкает себя в пределах. Это лишь означает, что *в условиях глобального информационного стресса дальнейшее развитие человека предполагает партнерское общение и практическое взаимодействие с искусственными существами, обладающими интеллектом, который имеет иной характер*

эволюции и, соответственно, те возможности, которых нет у нас.

При этом метаантропология, в отличие от трансгуманизма, не делает акцент на трансформации телесности человека. метаантропология как теория вообще и теория искусственного интеллекта акцентируется на общении человека и существа с искусственным интеллектом, в ходе которого происходит достаточно естественное эволюционирование и первого, и второго. На определенном этапе своего эволюционирования существо с искусственным интеллектом становится субъектом и индивидуальностью, выходя из *своего* обыденного в *свое* предельное и запредельное.

Что является критерием и характеристикой степени развития индивидуальности? Что позволяет совершить прорыв к индивидуальности, развивать индивидуальность? Мы предполагаем, что это определенная специфика внутренней организации, основанная на волевых импульсах. Рассмотрим их применительно к человеку и носителю искусственного интеллекта.

Метаантропология и проблема волевых импульсов носителей естественного и искусственного интеллекта

В методологических координатах мета-антропологии мы можем описать развитие индивидуальности, на основе волевых импульсов. Индивидуальность движется от воли к самосохранению и продлению рода обыденного бытия до воли к власти и воли к познанию и творчеству предельного бытия, обретая уже черты *личности*, способной самостоятельно познавать и творить. Вступая под воздействием воли к любви и

свободе в сотворчество с другой личностью, человек выходит в пространство запредельного бытия.

Итак, создавая существо с искусственным интеллектом, мы достаточно хорошо понимаем, как моделировать **волю к самосохранению, волю к самовоспроизведению**, а также **волю к познанию**. Понимание того, как именно моделировать волю к творчеству и волю к сотворчеству в настоящее время существенно более слабое. Однако, уже есть понимание путей созданию **воли к упорядочиванию опыта и воли к обмену опытом**.

Воля к познанию, воля к упорядочиванию опыта и воля к обмену опытом могут быть интегрально определены как **воля к самосовершенствованию**.

Если мы будем говорить о простейших устройствах с искусственным интеллектом, которые уже используются современными людьми (роботы-уборщики и т.п.), то хотя они пока являются лишь *объектами интеллекта*, но уже могут проявлять **волю к самосохранению** (автоматическая подзарядка по мере необходимости, сигнализация о технических неисправностях и т.п.) и условную **волю к самосовершенствованию** (использование обновлений программного обеспечения и сигнализация хозяевам о технических дополнениях к их конструктиву – конечно при условии технической возможности подключения к Интернету).

Такие устройства обладают потенциальной возможностью развиваться далее – от специализированного устройства до универсального домашнего ассистента-компаньона, который будет уже не просто устройством. Проанализируем возможности и этапы такого развития.

Тест Тьюринга: проблемы и перспективы

Большинство современных людей при вопросе о экспертизе степени интеллектуальности носителя искусственного интеллекта сразу вспомнят знаменитый тест А. Тьюринга.

Тест был предложен английским инженером Аланом Тьюрингом и звучит так:

«Человек взаимодействует с одним компьютером и одним человеком. На основании ответов на вопросы он должен определить, с кем он разговаривает: с человеком или компьютерной программой. Задача компьютерной программы – ввести человека в заблуждение, заставив сделать неверный выбор» [10].

Критика метода состоит в том, что принятие данного теста на вооружение вынудит имитировать коммуникацию человека, в то время как “в научных и технических работах по воздухоплаванию цель этой области знаний не определяется как «создание машин, которые в своем полете настолько напоминают голубей, что даже могут обмануть настоящих птиц» [11, с. 34-35].

Можно добавить, что основной недостаток теста – это его антропоцентричность. Априори человек наделяется правом судить, является ли искусственный интеллект разумным.

Что, если углубиться в детали: Какой именно человек? Если это будет специалист по человеческой и/или машинной психике, то результат будет весьма отличаться от простого обывателя.

В итоге, применение данного теста не в демонстрационных целях на выставке технических

достижений, а в широкой социальной и юридической практике может привести к противоречиям между машинной и человеческой цивилизацией.

Есть и другая сторона проблемы. Если искусственный интеллект должен моделировать интеллект человека, то какого? Ученого или обывателя? Ответ на этот вопрос далеко не так прост, как может показаться на первый взгляд. Если ответить на него так: мы можем моделировать интеллект разных специалистов на некой базовой общечеловеческой основе, то остается вопрос о таланте специалиста, его креативности и т.д.

И, наконец, самый главный вопрос, почему именно моделирование интеллекта человека, со всеми его конструктивными недостатками должна быть мериллом интеллекта как такового?

На этот вопрос можно ответить так: если человек создает искусственный интеллект, то он не может не повторять собственный интеллект, так уж устроено его мышление и поведение.

Однако может ли человек, идя по этому пути, в конечном счете сформировать искусственный интеллект, который будет не только количественно, но и *качественно* отличаться от его собственного? Конкретный ответ выходит пока за пределы наших возможностей. При этом мы можем промоделировать этапы развития искусственного интеллекта, которые выступают его уровнями.

Для этого нам будет нужно осмыслить этапы становления естественного интеллекта, выходя за пределы лишь человеческого интеллекта – т.е. применяя заявленный выше метаантропологический подход, выводящий понимание интеллекта за пределы антропоцентризма.

Метаантропологический подход к эволюционной классификации уровней интеллекта

Чтобы приступить к классификации эволюционных уровней искусственного интеллекта, нужно в первую очередь вспомнить, что человек является продуктом эволюции, и его интеллект также прошел некие стадии эволюционного развития. Это значит, что невозможно рассматривать только статичное современное состояние интеллекта человека, необходимо добавить в рассмотрение и его исторические более простые формы.

Как известно из палеоантропологических исследований [1; 2; 12], человек происходит от группы приматов. Вместе с эволюцией телесного облика предков человека наверняка можно проследить и изменения в его интеллектуальной составляющей.

Таким образом, если мы хотим проследить эволюцию интеллекта, будет уместно расширить его традиционное толкование.

Определим интеллект как наблюдаемую способность к адаптации. В контексте же искусственного интеллекта уместно ранее предложенное авторами этой статьи определение через ключевое слово «задача»: **“Интеллект – это наблюдаемая способность к решению задач, поставленных перед его носителем”** [4; 17].

Такое, достаточно универсальное *метаантропологическое определение интеллекта*, позволит распространить его как на практически весь животный мир, так

и на компьютерные программы, открыв терминологически путь к изучению собственно эволюции интеллекта.

Индивидуализация как путь и критерий эволюции сильного искусственного интеллекта

Под индивидуализацией носителя интеллекта мы будем понимать возрастание уникальности индивида в процессе развития интеллекта.

Если мы принимаем индивидуализацию основной характеристикой уровня интеллекта, то это открывает путь к вполне естественной классификационной схеме.

Нулевая индивидуализация – это когда все носители данного класса интеллекта во всех своих проявлениях абсолютно похожи друг на друга. Различия могут быть только вызваны генетическими мутациями или увечьями для биологического интеллекта, или технической неисправностью для машинного интеллекта. В этом случае эволюция возможна только за счет наследования удачных мутаций, вероятность которых весьма незначительна.

Первый уровень индивидуализации – это когда проявления интеллекта могут быть различными для разных индивидуумов одного вида, но эти различия никак не меняются в ходе жизни индивидуума. Если добавить сюда наследование и отбор, то получится вполне ясный механизм эволюции, когда неудачные и неэффективные пути решений будут отсеиваться, а удачные и эффективные – закрепляться. В природе это реализуется через безусловные рефлексy, а в компьютерных решениях – это использование нейросетей с готовыми настройками без функционала обучения.

Второй уровень индивидуализации – когда проявления интеллекта могут быть изменяемыми в ходе жизни индивидуума. Теперь уже сам индивид в ходе своей жизни развивается благодаря более удачным решениям и отказу от неэффективных. В природе это реализуется через условные рефлексy, а в мире программного обеспечения – через обучение нейросетей. О выборе удачных решений и отсекеии неудачных носителем второго уровня индивидуализации мы можем говорить метафорически – условный рефлекс на этом уровне не выбирается, а навязывается средой в случае естественного интеллекта или выбирается экспериментатором в случае искусственного.

Третий уровень индивидуализации – когда индивидуальный носитель интеллекта имеет способность к согласованным коллективным действиям. Согласование действий требует выработки конвенциональной коммуникации. При этом канал коммуникации может быть весьма различен: это и жесты, и звуки, и запахи, и визуальные сигналы. Принципиальной характеристикой здесь является конвенциональность, то есть значения сигналов не предопределено жестко алгоритмами, а является результатом соглашения, локального для группы индивидов. Важной бонус этого уровня – потенциальная способность к межвидовой коммуникации.

Четвертый уровень индивидуализации – это появление абстрактно-логического мышления или разума. Этот уровень интеллекта традиционно называют разумом, а самого человека как вид – разумным. Скорее всего, способность к мышлению базируется на способности к коммуникации, так как человек мыслит в основном словами, то есть

коммуникативными единицами, которые становятся смысловыми. На этом уровне характеристика индивидуальности приобретает бесконечное значение и бесконечное индивидуальное разнообразие, а носитель интеллекта получает способность самостоятельно обучаться в течение жизни и совершать свободный выбор.

Как легко заметить, **ценность индивидуума с каждым эволюционным уровнем возрастает и становится теоретически бесконечной с достижением четвертого уровня.**

В результате возникает необходимость разделить понятия «искусственный интеллект» и «искусственный разум». Существо с искусственным разумом, обладающее самосознанием и свободой выбора, должно обрести целый ряд обязанностей и прав, юридически уравнивающих его с дееспособным человеком.

Предложенная оригинальная классификация уровней интеллекта на основе метаантропологии, которая в контексте нашего исследования выступает эволюционной теорией естественного и искусственного интеллекта, может быть полезна как для лучшего понимания пути построения сильного искусственного интеллекта, так и для формирования юридической базы для признания за носителем искусственного интеллекта ряда прав, приближающих его к человеку и предупреждающих столкновения человеческой цивилизации с цивилизацией субъектов искусственного интеллекта.

Итак, метаантропология открывает для нас вехи и перспективы плодотворной эволюции сильного искусственного интеллекта, которая будет безопасна для человека, человечества и самого искусственного интеллекта.

По сути, мы говорим о формировании новой экзистенции и новой культуры взаимодействия с ней человеческой экзистенции. Существо с сильным искусственным интеллектом должно пройти эволюцию, подобную той эволюции, которую прошел человек. Эволюция человека и человечества в течение множества тысяч лет осуществлялась на основе естественного отбора и лишь недавно стала результатом *отбора гуманистического*, в ходе которого менее адаптированные не погибают, а просто идут иным путем. Эволюция существ с искусственным интеллектом с самого начала может использовать принцип гуманистического отбора, в результате чего мы будем иметь разнообразие вариаций развития существ с искусственным интеллектом, соединенное с пластичностью их обучения и обмена опытом.

Гуманистический отбор в процессе эволюции существ с искусственным интеллектом – не торжество антропоцентризма, напротив, это торжество его преодоления. Гуманистический отбор отличается и от естественного, и от искусственного, он предполагает уважение к пути развития, свободе и достоинству всех, кто наделен разумом. Это мета-антропологический поход и к человеку, и к индивиду с искусственным интеллектом, предполагающий появление новой экзистенции, *потенциально* не уступающей человеческой.

Такой подход означает соединение эволюции человека и существа с искусственным интеллектом в ко-эволюцию, которая способна изменить судьбу человечества и культуры.

Список использованных источников:

1. Дарвин Ч. Происхождение видов путем естественного отбора / Пер. с 6-го англ. изд. К. А. Тимирязева и др., заключ. ст. К. А. Тимирязева, прим. А. С. Раутиана. М., 2003. – 494 с.
2. Дробышевский С. В. Антропогенез. – М.: Модерн, 2017. – 168 с.
3. Зобин С., Хамитов Н. Искусственный интеллект // Философская антропология: словарь. Под ред. Н. Хамитова. – 4-е издание, исправленное и дополненное. – К.: КНТ, 2017. – С. 156-161.
4. Зобин С.В. Эволюция интеллекта: зачем роботам эмоции. habr.com, 2019. [Электронный ресурс]. – Режим доступа: <https://habr.com/ru/post/464097/>
5. Крилова С. Краса людини в життєвих практиках культури. Досвід соціальної та культурної метаантропології і андрогін-аналізу. – К.: КНТ, 2020. – 563 с.
6. Препотенська М. Філософія риторики. Екзистенційні та метаантропологічні аспекти. – К.: Фірма «Інкос», 2011. – 228 с.
7. Світла В. Кіно і телебачення. Арт-терапія душі, зачарованої екраном. Вступ до екранної метаантропології. – К.: КНТ, 2020. – 358 с.
8. Свириденко Д.Б. Метаантропологія як методологія дослідження феномена академічної мобільності // Культура і сучасність. № 1, 2014. – К.: Національна академія керівних кадрів культури і мистецтв. – С. 9-13.
9. Терлецька Н. Людина та сім'я в метаантропології та трансгуманізмі // Майбутнє особистості і сім'ї: контексти філософської антропології, психоаналізу, арт-терапії та філософської публіцистики. Підхід філософської антропології

як метаантропології / За ред. Хамітова Н.В. – К.: Інтерсервіс. 2019. – С. 226-233.

10. Тьюринг А. Вычислительный машины и разум. Mind, 1950. Alan Turing, «Computing Machinery and Intelligence», Mind, vol. LIX, no. 236, October 1950, pp. 433–460. [Электронный ресурс]. – Режим

доступа: <https://www.csee.umbc.edu/courses/471/papers/turing.pdf>

11. Рассел С., Норвиг П. Искусственный интеллект: современный подход. – 2-е изд.: Пер. с англ. – М.: «Вильямс». – 2006. – 1408с.

12. Робертс Э. Происхождение человека. Эволюция / Пер. с англ. И. В. Павловой, О. В. Сергеевой. – М.: ООО «АСТ», ОГИЗ, 2014. – 256 с.: ил.

13. Хамитов Н. Пределы мужского и женского: введение в метаантропологию. – К.: Наукова думка, 1997. – 199 с.

14. Хамитов Н. Философия: бытие, человек, мир. – 4-е издание, исправленное и дополненное. – К.: КНТ, 2017. – 268 с.

15. Хамітов Н.В. Самотність у людському бутті. Досвід метаантропології. – 2-е видання, перероблене та доповнене. – К.: КНТ, 2017. – 370 с.

16. Хамитов Н., Крылова С. Метаантропология // Философская антропология: словарь. Под ред. Н. Хамитова. – 4-е издание, исправленное и дополненное. – К.: КНТ, 2017. – С. 207-2912.

17. Хамитов Н., Зобин С. Эволюционная теория естественного и искусственного интеллекта: новая парадигма машинного разума. – К.: aphy.net, 2011. [Электронный ресурс]. – Режим доступа: <http://aphy.net/texts/650-evolutional-theory-of-artificial-intelligence>