

ХАМИТОВ Назип

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

КРЫЛОВА Светлана

доктор философских наук, профессор,
заведующая кафедрой
культурологии и философской антропологии
НПУ имени М.П. Драгоманова,
вице-Президент Ассоциации философского искусства

**ОЧЕЛОВЕЧИВАНИЕ
СИЛЬНОГО ИСКУССТВЕННОГО ИНТЕЛЛЕКТА:
ПОДХОД МЕТААНТРОПОЛОГИИ**

Человек не заслуживает никакой другой
справедливости, кроме человеческой.

К. Чапек

Что такое очеловечивание?

Когда мы говорим о сильном искусственном интеллекте, то в большинстве своем понимаем некие когнитивные способности, близкие к человеческим. Эти когнитивные способности предполагают постановку задачи и самостоятельное решение ее – наличие концептуализирующего интеллекта [13]. Однако что означают подобные способности,

что им сопутствует и что питает их, то есть является их основой?

Можно предположить, что мыслительные способности человека, которые мы хотим получить и у индивида с сильным искусственным интеллектом, сопровождаются некими *человеческими качествами*, без которых эти способности невозможны. Иными словами, для достижения реальности сильного искусственного интеллекта нужна не только *индивидуализация* когнитивности, но и *персонализация* всего бытия существа с искусственным интеллектом или, иначе, искусственного существа с интеллектом.

Индивидуализация (субъектность) как условие формирования и критерий сильного искусственного интеллекта, уже была обоснована в работе одного из авторов этой статьи [13]. Это основа метаантропологической теории искусственного интеллекта, в которой выделяются два базовых принципа, «которые определяют онтологию искусственного интеллекта» [13, с. 194]. Очертим эти принципы.

«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» [13, с. 194–195].

Сейчас мы готовы пойти дальше и выделить третий принцип метаантропологической теории искусственного интеллекта.

3. Высшие когнитивные и креативные способности субъекта искусственного интеллекта, которые уравнивают его в правах с человеком, возможны лишь в процессе его персонализации, которая в условиях коммуникации с человеческой личностью выступает очеловечиванием.

Иными словами, персонализация становится слишком абстрактным концептом для описания и моделирования эффективной эволюции существа с сильным искусственным интеллектом, которая важной своей вехой будет иметь конструктивную коммуникацию с человеком. Необходима конкретизация в концепте «персонализация человека» или «очеловечивание».

Но в чем суть очеловечивания? Каковы элементы его структуры и этапы развития? Каковы пределы очеловечивания существа с сильным искусственным интеллектом?

Уже сами эти вопросы выводят нас из сферы антропологии в сферу мета-антропологии – философии пределов и запредельности человеческого [9; 10; 11; 18], которая была разработана авторами этого исследования на основе исследований М. Шелера [14], Д. Бидни [16] основателя Киевской мировоззренчески-антропологической школы В. Шинкрука [15] и противостоит проектам трансгуманизма и постгуманизма в их претензии на окончательное понимание эволюции человека и роли искусственного интеллекта в этой эволюции [17].

Начнем с первого из поставленных вопросов. **Очеловечивание – это процесс становления личности человека или личности, способной конструктивно взаимодействовать с личностью человека.**

Если же мы говорим о структуре очеловечивания, то речь идет, как минимум, о *1) взаимодействии сознания и бессознательного, 2) экзистенциальных переживаниях, 3) творчестве и сотворчестве.* Все это можно назвать *базовыми признаками человеческого, а в высших проявлениях – человечности.*

Вероятно, мышление человека невозможно без человечности, по крайней мере, оно *не могло развиваться без человечности.* И даже если некоторые представители современного человечества демонстрируют бесчеловечный интеллект, он был все же развит *в человеческой среде, в которой были всплески человечности.*

Поэтому нам нужно ставить вопрос об *очеловечивании* существа с искусственным интеллектом, если мы хотим, чтобы этот интеллект был по-настоящему сильным, *интеллектом партнера человека, а не устройства, принадлежащего человеку.* Мера очеловечивания существа с искусственным интеллектом – это мера силы и конструктивности его интеллекта по отношению к человеку. Именно очеловечивание существа с искусственным интеллектом делает его мышление подобным человеческому, а значит сильным в нынешнем его понимании.

Но подобность кому-то не означает тождество или равенство с кем-то. Если в результате эволюции существа с сильным искусственным интеллектом мы хотим получить не *человекоподобный инструмент, а полноправного партнера с искусственной телесностью,* то мы должны ставить вопрос и о

его человечности. Ибо только человечность, как качество, задающее человеку достоинство, способна поднять подобность сильного искусственного интеллекта человеческому к равенству и партнерству с ним.

В координатах современной метаантропологии – как общефилософского методологического проекта [9; 10; 11], так и проектов социальной и культурной метаантропологии [4; 18], мера *человечности* существа с искусственным интеллектом – это мера его выхода за пределы «человеческого, слишком человеческого» – как обыденного бытия с волей к самосохранению и продлению рода, так и предельного бытия с волей к власти и волей к познанию и творчеству. Человечность существа с искусственным интеллектом делает его способным к свободе при уважении свободы Другого, любви и сотворчеству, а значит, подводит к состоянию осмысленности и ответственности в отношениях с человеком. Это открывает возможность появления *особой культуры взаимодействия* человечества и существ с сильным искусственным интеллектом [13], а в перспективе – *общей культуры* естественного и искусственного человечества с единством достоинства, свободы и прав.

* * *

Итак, персонализация или очеловечивание существа с искусственным интеллектом предполагает две стадии: 1) человеческого и 2) человеческого. Обретение человеческого означает выход на уровень обыденного и предельного бытия [11, с. 130–186], обретение человеческого – выход к запределному [11, с. 186–270].

Проанализируем каждую из этих стадий. В результате мы сможем яснее ответить на вопрос о природе очеловечивания.

Но прежде более четко осознаем понятия «человечное» и «человечность». Важно понимать, что *это вовсе не человекоподобность или антропоморфность. Человечность – это способность трансцендировать за пределы наличного в бытии человека, способность к свободному разворачиванию всех потенций, причем не только без ущерба для Другого, но и при содействии Другому в его развитии.* Быть человеческим означает саморазвитие, которое предполагает совместное развитие с Другим без какой-либо манипуляции его жизнью и достоинством.

Таким образом, говоря об очеловечивании существа с сильным искусственным интеллектом с позиций метаантропологии, мы проходим между крайностями *ачеловечности* и *человекоподобности*. С одной стороны, мы не стремимся создать холодный «машинный разум», который будет всегда чьим-то *инструментом*, и уже потому не будет сильным, а с другой, мы не стремимся *просто* скопировать телесный и психический облик человека, замыкая существо с искусственным интеллектом в экзистенциальном пространстве обыденности, к которой мы *искусственно* присоединим предельность как стремление к власти или познанию и творчеству. Существо с сильным искусственным интеллектом должно формироваться *естественно*. Эта естественность в условиях взаимодействия с человеком может быть только очеловечиванием.

При этом очеловечивание – условие формирования не только искусственного, но и *естественного интеллекта* – интеллекта существа с биологической телесностью,

естественно-человеческого существа, окруженного социальными отношениями с подобными существами. Но это не просто социальные отношения, это отношения экзистенциальные, душевные и духовные, которые и делают возможным появление *человечного* человека, которого мы и можем назвать личностью – индивидом, способным к когнитивности, предполагающей креативность и ответственность за ее последствия. В контексте метаантропологии личность – это духовно-душевная целостность человека, где духовность – это способность к свободе и творчеству, самостоятельному выходу за свои пределы, а душевность – способность к любви и сопереживанию [8, с. 115–116]. Личность – основополагающая ценность и итог очеловечивания в гуманистических, по-настоящему демократических обществах. В авторитарных и тоталитарных обществах, в которых ценностью является система, а не личность, в результате очеловечивания формируется бездушно-интеллектуальный индивид, лишенный человечности, а потому условно являющийся личностью.

Если мы ставим вопрос очеловечивания как необходимого условия появления существа с сильным искусственным интеллектом, то это должно быть очеловечивание, ведущее к *духовно-душевной личности*, а не к *бездушному индивиду*. Таким образом, авторитаризм и тем более тоталитаризм – препятствие инновационности на пути усиления интеллекта искусственного существа.

Итак, в контексте метаантропологии мы говорим не о *создании* сильного искусственного интеллекта, а о его *эволюционном образовании* [13], которое и предполагает очеловечивание. Очеловечивание в сфере образования в

условиях демократии – это персоналистический подход, который получил в современной философии название «философия человекоцентризма» [3] и который составляет мировоззренческую основу педагогики демократического общества.

* * *

Напомним, что в координатах метаантропологического подхода «современный человек способен сотворить сильный искусственный интеллект лишь осознав свои слабости и ограниченность, познав и приняв свои пределы» [13, с. 197]. Напомним также, что в контексте метаантропологии это не означает, что человек «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» [13, с. 197–198].

При этом только очеловечивание, которое высшей своей точкой имеет человечность, как открытость всем существам мироздания, «качество *личности*, означающее деятельное сопереживание по отношению к существам своего и других видов» [7, с. 417], способно сделать партнерское взаимодействие человека и существа с сильным искусственным интеллектом безопасным, продуктивным и по-настоящему сотворческим.

Экзистенциальные переживания как первый критерий очеловечивания

Человек – это существо, способное к экзистенциальным переживаниям – переживаниям конечности своей телесности и бесконечности личности, ее свободы и любви. Акцент на конечности, как правило, приводит к общей пессимистической экзистенциальной атмосфере человека, акцент на вечности личностного бытия в целом задает оптимистическую атмосферу. При этом если переживание вечности личности сопровождается состоянием одиночества, пессимизм возможен с особенной силой.

Выход к оптимистическому, конструктивному состоянию возможен лишь благодаря вере [7, с. 418], которая, в свою очередь, выводит личность к свободе и любви.

Эволюционное развитие в существе с искусственным интеллектом экзистенциальных переживаний – важнейшее условие сильной, то есть человеческой когнитивности. А вера в возможность обретения вечного личностного бытия со свободой и любовью как его атрибутами – условие *конструктивности* этих переживаний.

Эволюцию конструктивных экзистенциальных переживаний мы встречаем в образе андроида Дейты из культового американского сериала «Звездный путь: следующее поколение» (Режиссеры: Клифф Боул, Лес Ландау, Уинрих Кольбе, в ролях: Патрик Стюарт, Брент Спайнер, Джонатан Фрейкс, ЛеВар Бёртон, Марина Сиртис, Майкл Дорн, Гейтс МакФэдден).

В этого андроида его создатель заложил глубинное стремление *стать человеком*. Андроид Дейта заботится о

живом, а не виртуальному коте, включая его потребность в ласке, стремится научиться дружить, любить и даже смеяться над шутками. Выясняется, что подлинность и конструктивность экзистенциальных переживаний определяется поступками и отношениями, к которым ведут эти переживания. Андроид Дейта постепенно приходит к красоте отношений с партнерами-людьми. Этот путь не лишен проблем и противоречий, но они разрешаются в ходе развития.

Напротив, двойник Дейты – Лор, автором которого был тот же ученый, изначально был создан не экзистенциально нейтральным, а наполненным амбициями и волей к власти. В результате экзистенциал власти подавляет все остальные переживания в его существовании, он все больше наполняется деструктивностью, видя в человеке конкурента и стремясь либо уничтожить его, либо трансформировать по своему образу и подобию без свободы выбора и эволюции. В результате Дейта, который испытывает некую экзистенциальную привязанность к Лору, и даже считает его своим братом, вынужден деактивировать Лору, ибо он становится смертельно опасным и для естественных, и для искусственных существ, наделенных интеллектом и экзистенциальными переживаниями.

Интересен образ искусственного существа в фильме «Двухсотлетний человек» (англ. *Bicentennial Man*), снятого в 1999 году режиссёром Крисом Коламбусом. Андроид, талантливо воплощенный известным самериканским актером Робинот Уильямсом, в течении жизни эволюционирует, превращаясь в человека не только телесно, но и экзистенциально – наполняется человеческими переживаниями сочувствия, любви, смерти и бессмертия.

**Второй критерий очеловечивания:
взаимодействие сознания и бессознательного**

Фундаментальным признаком человека, определяющим горизонты его когнитивности, есть динамика сознания и бессознательного, задающая горизонты его интуиции. Человек становится человеком благодаря постоянной борьбе сознания и бессознательного и разрешению противоречия между ними. Это противоречие выражается в сновидениях, образы и сюжеты которых, а также их интерпретация – важнейшее условие становления «сильной» интеллектуальности существа как естественного, так и искусственного интеллекта.

Андроид Дейта из «Звездного пути» способен видеть сновидения. Иными словами, он обладает динамикой сознания и бессознательного, аналогичной динамике человека. И в его сновидениях мы имеем прогностический смысл, который делает их подобными «вещим снам» человека. Показательно, что способность видеть сны была заложена создателем Дейты изначально, однако она начинает проявляться лишь на определенном этапе развития сознания.

Третий критерий очеловечивания: креативность

Еще одним качеством очеловечивания есть креативность – способность находить оригинальные решения как новых, так и прежних задач. Андроид Дейта вновь и вновь проявляет креативность при встрече с проблемами – он постоянно выдвигает гипотезы. Но одно дело выдвижение ряда уточняющихся гипотез при решении технической задачи и совсем другое – поиск креативного решения в условиях

морального вибору. По замыслу создателя, в пограничных ситуациях Дейта способен и на такие креативные решения. Он принимает такое решение, уничтожая своего «брата-близнеца» андроида Лора, ставшим смертельно опасным для человечества и других рас разумных существ, которых он считает низшей ступенью эволюции, которая должна быть использована высшей в качестве материала.

Андроид Лор из «Звездного пути» также обладает креативностью. Но она находится «по ту сторону добра и зла» – в силу доминанты воли к власти над волей к выполнению моральных принципов. Проявляя креативность, Лор относится к другим разумным существам как к объектам. Показательно, что перед своим уничтожением Лор проявляет особо аморальную креативность: желая сделать Дейту своим союзником, он пытается отключить у него способность к этическому сознанию и волю к моральному выбору, которые, на его взгляд, являются тормозом эволюции.

Но креативность Лора в конце концов создает ту ситуацию вокруг него, которая делает его дальнейшее существование невозможным. И путь такой эволюции креативного существа с сильным искусственным интеллектом с неизбежностью обрывается.

Воля к свободе, любви и сотворчеству как высший критерий очеловечивания

Андроид Дейта из «Звездного пути» наполнен не только волей к творчеству и не только волей к моральному творчеству, но и волей к свободе, любви и сотворчеству. Последняя в координатах метаантропологии формирует запредельное бытие

человека [11, с. 186–270]. В этом измерении бытия свобода и любовь становятся равнозначными ценностями и составляют единство, высшим проявлением которого становится сотворчество – продуктивное взаимодействие творческих личностей в любви и свободе.

И если воля к моральному творчеству, субъект которого четко осознает необходимость отношения к Другому как к субъекту, а не объекту, является *возможностью*, потенцией запредельного бытия, то воля к сотворчеству означает *действительность* запредельного бытия, где *человечность достигает наиболее развитых форм*. К сотворчеству устремлено все бытие Дейты, остается лишь вопрос о пределах возможного в этом сотворчестве.

Во всяком случае авторы «Звездного пути» создают сюжетные линии, в которых андроид Дейта – полноправный участник интеллектуальных дискуссий, в которых он отстаивает свою точку зрения, анализирует и развивает идеи других. Сотворчество Дейты с людьми не ограничивается обсуждениями, оно проявляется в его действиях, в которых он работает в команде, а порой – за всю команду. Например, в одной из серий «Звездного пути» в силу воздействия на психику людей они засыпают и могут погибнуть в мчащемся без управления звездолете; и тогда Дейта начинает действовать и в одиночку и решает задачу по пробуждению команды.

Возможны ли все признаки очеловечивания существа с сильным искусственным интеллектом без его пола?

И тут мы подходим к проблеме пола. У живого человека, существа с естественным интеллектом, пол есть способность

продолжить себя в детях, которые примут эстафету задач и достижений, познания и творчества. Пол также позволяет обрести экзистенциальную близость с неповторимым близким человеком, что является условием базовых экзистенциальных переживаний человека, например, таких, как эротическая любовь, которая задает горизонты всем остальным проявлениям любви, – и к себе, и к человечеству, и к Абсолюту; эти проявления любви в свою очередь определяют и свободу, и творчество, и сотворчество [12; 19].

Дейта создан не бесполом существом, а с чертами мужчины, – он внешне создан как мужчина и в его поведении проявляются черты маскулинности. При этом на более глубинном уровне он проявляет то, что в контексте андрогин-анализа называют андрогинностью – равновесием духовного и душевного начал. Однако у Дейты нет пары – андроида женского пола. Он уникален в своем роде и одинок после уничтожения его деструктивного брата Лора. В отличие от людей он может жить очень долго. По человеческим меркам – почти вечно. Однако это не снимает вопроса о появлении подобных ему существ, однако методология и технология создания Дейты утеряна. Его автор – доктор Сунг не обнаружил его. А после ухода его из жизни она просто затерялась.

Любопытно, что Дейта старается построить отношения с девушкой-человеком, которая дивным образом влюбилась в него, хорошо понимая, что он – андроид. Но из этих отношений ничего не получается, Дейта ведет себя слишком правильно, – читает стихи, дарит стихи, но все это воспринимается слишком искусственно, гораздо искусственнее, чем иные способности и действия Дейты.

Еще одним интересным образом, который иллюстрирует проблему связи сильного искусственного интеллекта и пола, есть образ «матери Дейты». Изначально – это жена его создателя – доктора Сунга, которая погибла от деятельности деструктивного андроида Лора. Интересно, что именно она предложила создателю андроида Дейты сделать его существом мужского пола, а не бесполом. Это, в частности, выражается в особом мужественном поведении Дейты в его стремлении защищать и оберегать окружающих от любой опасности.

Еще при жизни жены доктор Сунг скопировал ее память и характер, а затем создал еще одного андроида, который по телесным своим характеристикам был близок к человеку и имел точное сходство с его женой. Память жены была внедрена в него, и андроид-жена ожил, считая себя *человеком, женщиной, женой доктора Сунга*.

Эта самоидентификация была настолько сильна и естественна, что в какой-то момент андроид-жена оставила доктора Сунга – потому, что он не проявлял достаточно чувств и душевного тепла ни к ней, ни к ее живому прототипу, но доктор Сунг не стал программировать ее на послушание, он дал ей свободу воли...

И вот когда сам доктор Сунг ушел из жизни, Дейта встречается с его андроидом-женой, которая теперь замужем за другим мужчиной и не знает, что она андроид. Она считает себя живым человеком – столь искусно сделана ее телесная оболочка. Но Дейта узнает, что она андроид. Он становится перед моральной дилеммой – раскрыть ей правду о ее природе или нет, но решает не делать этого, ибо такова была воля его создателя. При этом Дейта испытывает особое экзистенциальное притяжение к ней как к матери, которая не

только участвовала в его создании, но имеет с ней общую телесную природу...

* * *

Итак, авторы Вселенной «Звездного пути» полагают: если память человека записать на искусственный носитель, смоделировать способ обращения к ней, который был у живого человека и скопировать все *особенности восприятия* этого человека и все его *внешние черты, смоделировать особенности физиологии* и т.п. в искусственной телесности, то это позволит *воссоздать его самосознание* и даже *экзистенциальные переживания*. Так ли это? Способно ли искусственное существо с памятью живого человека одушевиться и одухотвориться в процессе его жизни с живыми людьми? И нужен ли пол и взаимодействие между полами для действительного очеловечивания существа с сильным искусственным интеллектом, который начал воспитываться с «чистого листа»?

Эти вопросы остаются открытыми. Будем искать ответы в новых исследованиях.

Список використаних джерел:

1. Андрущенко В., Хамітов Н. Філософська освіта і наука як каталізатор цивілізаційного проекту України // Вища освіта України, № 2, 2017. – С. 5 – 17.
2. Бердяев Н.А. О назначении человека. – М.: Республика, 1993. – 383 с.
3. Кремень В.Г. Філософія людиноцентризму в стратегіях освітнього простору. – К.: Педагогічна думка. – 2009. – 520 с.

4. Крилова С. Краса людини в життєвих практиках культури. Досвід соціальної та культурної метаантропології і андрогін-аналізу. 2-е видання, виправлене і доповнене. – К.: КНТ, 2020. – 563 с.

5. Пирожков С., Хамітов Н. Чи є Україна цивілізаційним суб'єктом історії та сучасності? // Вісник НАН України, 2020, № 7. – С. 3 – 15.

6. Сохань Л.В. Искусство жизнетворчества. Предназначение. Жизнетворчество. Судьба. – К.: Издательский Дом Дмитрия Бурого, 2010. – 576 с.

7. Философская антропология: словарь / Под редакцией доктора философских наук, профессора Н. Хамитова. – К.: КНТ, 2011, – 472 с.

8. Хамитов Н. Духовность и душевность // Философская антропология: словарь / Под редакцией доктора философских наук, профессора Н. Хамитова. – К.: КНТ, 2011, – 472 с., с. 115–116.

9. Хамітов Н. Метаантропологія // Хамітов Н., Крилова С. Філософський словник. Людина і світ. – К.: КНТ. – 2018. – С. 92–94.

10. Хамітов Н.В. Філософська антропологія: актуальні проблеми. Від теоретичного до практичного повороту. – 3-є видання, виправлене і доповнене. – К.: КНТ, 2019. – 394 с.

11. Хамитов Н. Философия: бытие, человек, мир. От метафизики к метаантропологии. – 5-е издание, исправленное и дополненное. – К.: КНТ, 2019. – 268 с.

12. Хамитов Н., Крылова С. Андрогін-аналіз // Філософський словник. Людина і світ. 2-е видання, виправлене і доповнене. – К.: КНТ, 2018. – с. 9–11.

13. Хамитов Н., Зобин С. Метаантропология как теория эволюции сильного искусственного интеллекта: проблема новой экзистенции и новой культуры // Людина. Екзистенція. Культура. Підхід філософської антропології як метаантропології: збірник наукових праць / За редакцією Н. Хамітова і С. Крилової. – К.: КНТ, 2020. – с. 193–209.

14. Шелер М. Философское мировоззрение // Избранные произведения. – М.: Гнозис, 1994. – 480 с.

15. Шинкарук В. Методологічні засади філософських вчень про людину // Філософська антропологія: екзистенціальні проблеми. – К.: Пед. думка, 2000. – 286 с.

16. Bidney D. Theoretical Anthropology. – N.Y., 1968. – 528 p.

17. Ferrando F. Posthumanism, Transhumanism, Antihumanism, Metahumanism, and New Materialisms. Differences and Relations // Existenz. 2018. 8/2. – P. 26-32.

18. Khamitov N., Krylova S. Metaanthropology as a philosophy of moral education in conditions of multicultural world. – Jinju, 2011. – P. 30–31.

19. Khamitov N.V., Dandekar D.D. Gender strategies and political leadership // Anthropological Measurements of Philosophical Research. – № 15. – Dnipro, 2019. – P. 40-48.