

## СОЦИАЛЬНАЯ ФИЛОСОФИЯ И ФИЛОСОФСКАЯ АНТРОПОЛОГИЯ

Научная статья

УДК 101.3

doi: 10.17223/1998863X/87/9

### МАНИПУЛЯЦИОННАЯ ЕМКОСТЬ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Евгения Викторовна Авдеенко

*Воронежский государственный технический университет, Воронеж, Россия,  
jjaane@yandex.ru*

**Аннотация.** В статье рассматривается манипуляционная емкость искусственного интеллекта в сравнении с естественным. Проводится анализ формирования психики и человеческого сознания, искусственного и естественного мышления. Предполагается, что основное отличие ЕИ и ИИ состоит в природе их мотивации. ИИ лишен инстинкта самосохранения, следовательно, имманентной потребности действовать, с одной стороны, и ограничений активности – с другой. Отсутствие инстинкта самосохранения у ИИ повышает его манипуляционную емкость и как объекта, и как инструмента этого процесса.

**Ключевые слова:** сознание, искусственный интеллект, манипуляция, инстинкт самосохранения, деятельность

**Для цитирования:** Авдеенко Е.В. Манипуляционная емкость искусственного интеллекта // Вестник Томского государственного университета. Философия. Социология. Политология. 2025. № 87. С. 91–103. doi: 10.17223/1998863X/87/9

## SOCIAL PHILOSOPHY AND PHILOSOPHY OF HUMANITY

Original article

### MANIPULATIVE CAPACITY OF ARTIFICIAL INTELLIGENCE

Evgeniya V. Avdeenko

*Voronezh State Technical University, Voronezh, Russia Federation, jjaane@yandex.ru*

**Abstract.** This article examines the manipulative capacity of artificial intelligence (AI) in comparison with natural intelligence (NI). To this end, the study analyzes the development of the psyche in phylogenesis and its highest form—human consciousness. The specific features of human thinking are delineated in contrast to animal cognition, with the primary distinguishing characteristic identified as the human ability to operate at an abstract level. A comparative analysis of artificial and natural thinking is conducted, leading to the conclusion that the difference in their cognitive processes is more quantitative than qualitative. A further distinction lies in the substrate of thought: NI is realized through electrical impulses within a neuro-humoral environment, whereas AI functions within a

silicon-based environment. However, the fundamental difference between NI and AI lies in the nature of their motivation. AI's motivation is assigned by its creator-programmer, while the source of motivation for NI is the self-preservation instinct. For NI, this instinct serves as both an impulse for development and activity, and a regulator that imposes limits. In contrast, AI lacks a self-preservation instinct and is thus devoid of both an inherent drive to act and any intrinsic constraints. This absence of self-preservation in AI infinitely expands its manipulative capacity, both as an object and as an instrument of manipulation. Figuratively, AI operates on the principle of "I see the goal, I see no obstacles". For NI, the self-preservation instinct acts as a determinant of control and self-restraint, forcing a correlation between the value of a goal and the potential losses incurred in its pursuit. Since self-preservation is not a "built-in option" for AI, it can pursue a goal without accounting for losses, even to the point of self-destruction. The novelty of this research lies in its differentiation of the determinants of NI and AI specifically within the context of manipulative capacity.

**Keywords:** consciousness, artificial intelligence, manipulation, self-preservation instinct, motivation, activity

**For citation:** Avdeenko, E.V. (2025) Manipulative capacity of artificial intelligence. *Vestnik Tomskogo gosudarstvennogo universiteta. Filosofiya. Sotsiologiya. Politologiya – Tomsk State University Journal of Philosophy, Sociology and Political Science*. 87. pp. 91–103. (In Russian). doi: 10.17223/1998863X/87/9

## Формирование сознания

Человеческое сознание возникает в процессе интериоризации форм общественного сознания. Индивид не обладает человеческим сознанием «по рождению». Доказательством тому служат примеры «детей-маугли»: если человеческий детеныш не находится в человеческой социальной среде, он не обретает человеческого сознания.

Высшим этапом развития психики является человеческое Сознание [1].

Сознание человека отличается от психики животных способностью к абстрактному мышлению. Это возможность оперировать неопредмеченными понятиями. Крайнее выражение абстракции, на наш взгляд, представлено в математике. Цифра – это ничто, с одной стороны, но она имеет огромное значение в сознании человека и человечества. Ключ отличия человеческого от животного сознания состоит именно в возможности выйти на уровень абстрактного мышления. Животное способно к интеллектуальному поведению, научению, формированию условных рефлексов. Но оно не обладает способностью опосредовать свою деятельность и мышление второй сигнальной системой [2]. Собака может научиться считать, но считать лишь предметы. Она может запомнить цифры и выстраивать свое поведение в соответствии с их предметным или форменным образом. Но ей не доступен их абстрактный смысл, лишь форма образа, который их опредмечивает. Например, если в Т-образном лабиринте проход, обозначенный знаком 4, сулит животному удар током, а 5 – пищу, оно усвоит выгодное для него направление, но это связано только с тем, что оно воспринимает форму написания как предмет.

Интересно, что обретение человеческой психики доступно для человеческого индивида не безвременно. Если человеческий ребенок не оказывается в человеческом социуме до 5–7 лет, он не только не научится говорить, он даже не сможет освоить прямохождение как естественный для себя способ перемещения.

Абстрактный уровень мышления открывает для человека следующие возможности.

Во-первых, способность к теоретическому обучению. Кошке для обучения котенка охоте необходим конкретный воробей, а человек может обучить другого человека, не оперируя предметами непосредственно.

Во-вторых, абстрактное мышление позволяет нам организовывать сложную совместную деятельность. Да, стайные животные могут совместно охотиться. Но для этого им нужен предмет охоты в непосредственной данности. Волки не могут, как футболисты, заблаговременно обсудить, что будет делать каждый участник команды, если «заяц бежит с левого фланга с определенной скоростью».

В-третьих, абстрактное мышление является средством планирования. Может возникать иллюзия, что и животное может регулировать свою деятельность предполагаемыми перспективами, например, терпеть и не совершать дефекации в неположенном месте, опасаясь наказания. Но это не планирование, а лишь оперантное научение [3].

И самое главное – животное не может опосредовать свою деятельность духовными или идеологическими потребностями, реализация которых всецело находится в абстрактном поле.

Таким образом, именно абстрактное мышление позволяет человеку выйти за рамки предметного опыта. Обучаться и действовать не только вследствие оперантного и тем более респондентного научения [3], но опосредуя свою деятельность второй сигнальной системой [2].

В филогенезе человеческая психика и мышление проходят стадии от сенсорного, инстинктивного поведения, к научению, интеллектуальному мышлению и абстрактному [1].

Сознание развивается постепенно на основе исключительно телесного опыта и только в постоянной коммуникации с другими субъектами [4, 5].

Искусственный интеллект сразу оказывается в рамках абстрактного пространства. Он в принципе «незнаком» с предметным миром. Тканью естественного интеллекта являются вполне материальные белковые связи и электрические импульсы. Тканью искусственного интеллекта является все то же электричество и кремниевые связи. Ментальной сущностью ЕИ являются образы и знаки. Для ИИ это лишь цифры – абсолютная абстракция. Еще более утопичным кажется утверждение, что это лишь две цифры: ноль и единица – онтология ИИ заключена в бесконечной последовательности этих двух знаков.

Обратимся к объектно-ориентированной онтологии [6].

С точки зрения классической субъектной онтологии все есть лишь в соотношении, понимании или даже непонимании человеческого субъекта. Объектно-ориентированная онтология обращает внимание исследователей на то, что сущностью, а главное – волей, могут обладать объекты вне аффективной связи с человеческим субъектом. Классическая «дохайдеггеровская философия» не предполагала такого онтологического свойства объектов. Экзистенциальное хайдеггеровское «dasein» постулирует имманентное свойство человеческого сознания ощущать собственную «выдвинутость из объективной реальности», ощущать значимость собственного выбора не только для себя, но и для действительности. Крайним выражением уверенности в субъектной опосредованности бытия можно назвать «эффект бабочки» – всякое действие субъекта может совершить революционные события в объективном мире.

Хотелось бы обратить внимание на излишнюю амбициозность экзистенциалистов и гуманистов в приписывании «*dasein*» всякому человеческому субъекту. Большая часть представителей человечества пребывает в состоянии «*das man*», никоим образом не рефлексировав своей экзистенциальной сущности, ответственности и источников мотиваций. Прimitивно выражаясь – они просто «хотят», не задумываясь ни о причинах, ни о последствиях.

В вопросах детерминации бытия мы склонны придерживаться позиций объектно-ориентированной онтологии: объекты могут взаимодействовать и бытийствовать вне связи с человеческим субъектом. Мы не склонны к элиминации субъектности мышления лишь по признаку отсутствия в нем человеческого начала.

Мышление – это совокупность таких свойств сознания, как анализ и синтез, – мыслящее существо анализирует загруженные в сознание данные и синтезирует выводы, которые воплощаются в физических или ментальных паттернах. «Стратегия мозга состоит в постоянном построении «опережающих моделей», вероятных предвидений и использовании входящей информации для корректировки в случае необходимости» [7. С. 212].

Определение сознания, данное В.И. Лениным и до сих пор признаваемое психологами и философами актуальным, гласит: сознание – это высшая форма психического отражения объективной реальности. Человеческое сознание опосредовано интериоризацией форм общественного сознания, возникших и возникающих в процессе культурно-исторического развития. Если обратиться к этому и многим другим определениям психики и ее высшей формы развития – Сознанию, ничего специфически «человеческого» в нем нет. Да, так вышло, что на момент создания этих определений способность отражать культурно-историческую реальность была только у человека. Но на современном этапе развития мышления в его онтологическом и гносеологическом смысле оно стало присуще не только человеческому субъекту. Независимо от белковой или кремниевой основы в процессе мышления осуществляется анализ и синтез данных. Животные могут оперировать лишь предметно-чувственными данными. Искусственный интеллект – абстрактными, а человек и теми, и другими. Обратимся к вопросу, что есть чувственные данные? Существует иллюзия, что животные и человеческие ощущения принципиально отличаются от восприятия искусственных сенсоров. Однако на сегодняшний момент нет никаких данных функциональной диагностики, которые принципиально и качественно отличали бы систему человеческого декодирования и сенсорного распознавания ИИ. Информация, поступающая через сенсоры биологического существа, также декодируется нервной тканью. Мышление не есть модель мозга, «а модели предметной деятельности опосредования, в которой объединены: предмет (орудие) – орган – зона специализации мозга – предметное действие – знак, опосредующий действие» [8. С. 488].

Строго говоря, мышление человека и ИИ ничем не отличаются – оно есть анализ и синтез поступающей информации, в результате которого может осуществляться поведенческий или ментальный деятельностный выбор.

## **Источник активности Искусственной и Естественной Субъектности**

Всякому белковому организму, независимо от идеального или материального статуса его создателя, имманентно присуще стремление к самосохранению. Это мы можем наблюдать даже на примере самых простых форм существования живых организмов.

И именно стремление к самосохранению детерминирует развитие белкового организма и в онтогенетическом, и в филогенетическом контексте. «Вопрос выживания человека как биологического вида в ходе всей его эволюции был ключевым и стимулировал развитие его интеллекта» [9]. «Отсутствие тела у искусственного разума тормозит его развитие и даже исключает формирование «сильного» ИИ» [10. С. 5].

Искусственный интеллект, основу жизнедеятельности которого составляют небелковые соединения, лишен имманентного стремления к самосохранению. По крайней мере такие выводы мы можем сделать, наблюдая за «активностью» живой и неживой природы. Объекты неживой природы не обнаруживают какой-либо деятельности в направлении самосохранения. Хотя невозможно отрицать способность к осуществлению деятельности предметов и явлений неживой природы. Например, электрический разряд с онтологической точки зрения имеет деятельностьную основу – его можно оценить в категориях и процесса, и результата.

Как и в случае с молнией, мы не можем отрицать бытие и деятельность искусственного интеллекта. Молния – это лишь совокупность магнитно-физических факторов, меняющих тем не менее объективную реальность. Искусственный интеллект – это совокупность свойств анализа и синтеза, присущая специфическим формам организации неживой природы и обладающая способностью действовать. С точки зрения отечественной теории деятельности деятельность – это совокупность любых действий, которые направлены на достижение поставленных целей [11]. Именно деятельность формирует сознание [1]. Таким образом, с точки зрения способности к деятельности ИИ обладает и Сознанием, и Мышлением. Формально его деятельность опосредована определенными целями. Однако источник этих целей для ИИ на сегодняшний день всегда человек.

Критики объектно-ориентированной онтологии апеллируют к отсутствию Самосознания у мыслящего объекта. Еще Кант говорил: «То обстоятельство, что человек может обладать представлением о своем Я, бесконечно возвышает его над всеми другими существами, живущими на Земле. Благодаря этому он личность, и в силу единства сознания при всех изменениях, которые он может претерпевать, он одна и та же личность, т.е. существо, по своему положению и достоинству совершенно отличное от вещей» [12. С. 357].

И материалистическая, и гуманистическая, и экзистенциальная, и психоаналитическая, и все нео-... философские и психологические теории личности утверждают и доказывают тот факт, что личность – это проекция психики, основой которой, бесспорно, является белковая генетическая субстанция, но характеристики которой детерминируются средой и интериоризацией предметного и абстрактного ментального опыта. Самосознание – лишь свой-

ство отражения реальности, продуктом которого является восприятие мыслящего существа себя фигурой, выдвинутой из фона, в котором реализуется ее ментальная деятельность.

Объектно-ориентированная онтология преодолевает картезианский дуализм протяженной и мыслящей субстанции.

Чем более мы погружаемся в изучение особенностей сознания и мышления, тем меньше мы находим различий этих процессов в исполнении ЕИ и ИИ.

Принципиальное, качественное отличие мы можем наблюдать в источнике активности или деятельности. Всякая деятельность ЕИ опосредована имманентно присущей всем белковым существам потребностью в самосохранении. А у ИИ такая потребность не является имманентной. Она может быть задана программным кодом. И в этом смысле – создателем, т.е. человеком. Но в случае с белковыми существами такая потребность задана «создателем» в 100% случаев существования организма.

Бесспорно, эта потребность может реализовываться и блокироваться самими различными способами. Но именно и только она является источником какой бы то ни было деятельности белкового существа. «Тело, думается, тоже не обязательно должно иметь человеческий вид, это может быть даже какая-то область заинтересованно контролируемых сетей, но такие области должны быть сравнимы с другими областями, находящимися под контролем других искусственных интеллектуальных систем, и между этими системами должна быть заинтересованная коммуникация, включающая взаимные оценки, реакции одних на действия других» [13. С. 55].

Потребность в самосохранении – основной источник любой мотивации белковых организмов. А кремниевому-программному организму мы можем задать любую мотивацию.

В этом контексте принципиальное значение имеет этический аспект. Интериоризация форм общественного сознания формирует в личностной структуре Супер-эго – морально-этическую инстанцию. Этика для ЕИ выступает ограничителем: биологический организм может как угодно самосохраняться, развиваться и удовлетворяться, но если формы реализации этих процессов могут привести к гибели (физической или социальной), супер-эго блокирует соответствующую деятельность. Но она же и мотиватор – независимо от степени осознания нравственных принципов человек стремится стать лучшей версией себя [14, 15].

В случае с ИИ возникает гносеологический парадокс: с одной стороны, у него нет имманентной потребности в деятельности, с другой – нет ограничителей деятельности, детерминированных инстинктом самосохранения. Прimitивно выражаясь: он ничего не хочет делать, но может делать все, что угодно.

## **Манипуляционная конструкция**

«Манипуляция – это форма властных отношений посредством латентного, высокопрофессионального, технологического управления поведением через формирование в психических структурах человека определенных целей, установок, поведенческих паттернов и ценностей, где индивид учитывается как объект, а не как личность (в классическом определении личности

отечественной психологии)» [16. С. 586]. В этом определении подчеркивается тот факт, что, с точки зрения манипулятора, субъект онтологически является объектом. То есть для осуществления манипуляции субъектные свойства реципиента не значимы, он априори мыслится объектом.

Обратимся к представлению о том, что социальный мир, воплощенный в общественном сознании, есть совокупность дескриптивных и прескриптивных знаний. Дескрипция – знания о том, что есть, прескрипция – знания о том, что должно [17]. Принцип Юма (или Гильотина Юма) гласит: от суждений со связкой «есть» невозможно логически-непротиворечивым путем перейти к тому, что должно. То есть нет ничего сущего, «объективно» предписывающего нечто должное. Никакие факты реальности не могут стать обязательством для какого-либо поведения. Девид Юм высказывает это утверждение в «Трактате о человеческой природе» 1740 г. [18]. Затем этот принцип развивается и трактуется в исследованиях Пуанкаре и Поппера. На сегодняшний день не существует научного опровержения этого принципа.

Существует доктрина автономии морали, исходя из которой, морально этические нормы и не должны, и не могут исходить из сущего материально, потому как являются областью трансцендентного бытия субъекта.

Конструкты должного, как имманентного и трансцендентного, вступают в противоречие с принципом Юма. Имманентно должное становится как бы сущим в этом контексте. Однако, на наш взгляд, это и есть основа всякой манипуляции.

Манипуляция – это формирование ассоциативной связи между тем, что есть, и тем, что должно.

Здесь стоит учитывать, что манипуляция всегда имеет конечной целью изменение поведения. Не просто знания, а совершения определенных действий на основании этого знания. Без поведенческого компонента это не манипуляция, а идеология, или индоктринация, или даже просвещение.

Наше поведение детерминировано прескриптивным знанием. Прескрипция – это лишь предмет общественного договора. И в этом смысле это всегда манипуляция. Что бы ни было источником знания о том, как должно поступать, это никакого отношения не имеет к материальным заданностям.

Несмотря на все теории имманентной этики, для обоснования новых максим социального договора манипулятор каждый раз преодолевает принцип Юма. Выражаясь бытовым языком, он обманывает реципиента, выстраивая в его сознании ассоциативные абстрактные связи между тем, что есть, и тем, что должно. Например: где-то кем-то нарушаются права граждан, в связи с этим кто-то должен пойти и совершить какие-либо физические действия в отношении этих нарушителей. Таким образом, формируются ассоциативные причинно-следственные связи в абстрактном поле и закрепляются в общественном сознании, становясь затем основой для определенного поведения. Так инициировались Крестовые походы и Оранжевые революции. Так осуществляется массовая манипуляция. Интересно, что и всякая другая манипуляция происходит аналогичным образом. Эффективность манипуляции в принципе обеспечивает Система 1 [19]. Система 1 – эвристическое мышление, опосредованное эмоциями, образами и примитивными закономерностями. По сути, это мифологическое сознание. Чем отличается мифологическое

сознание от религиозного или секулярного? Наличием не веры, но знания. В вопросах веры субъект остается субъектом, т.е. не лишается права выбора: верить или не верить. Мифологическое сознание трансформирует субъект в объект – он не верит, он знает, т.е. задан.

Всякая манипуляция обращается к Системе 1 или мифологическому сознанию. Возникает вопрос – как же тогда работает манипуляция с людьми, склонными к постоянному использованию Системы 2 (рационального мышления) – лично или профессионально вовлеченными в интеллектуальную деятельность. Ответ – еще лучше. Ввиду многочисленных успешных опытов использования рационального аналитического мышления у профессионала формируется ощущение успешности своей интеллектуальной деятельности в принципе – доверия к ней. И в тот момент, когда манипулятор апеллирует к его Системе 1 (мифологическому сознанию), такой человек даже с большей верой принимает манипуляционные интроекты, так как привык доверять своему мышлению, не рефлексировав его дискриминанты. Человек «в белом халате» склонен верить другому «белому халату», даже если первый – лаборант физической кафедры, а второй рекламирует зубную пасту – так работает эвристическое мифологическое сознание.

В обществе постмодерна (и метамодерна, который наблюдают некоторые исследователи сегодня) основным провайдером между сущим и должным становятся симулякры [20] – образы, апеллирующие к сущему, но формирующие причинно-следственные связи от этого сущего к поведению физическому или ментальному – формированию мнений.

С точки зрения программирования симулякр – это простой алгоритм – последовательность образов и связанных с ним суждений.

Искусственный интеллект – это математическая модель, в которой заключено бесконечное количество «причинно-следственных» связей. Мы берем эти слова в кавычки, потому как на самом деле элементы этих связей не являются ни причинами, ни следствиями. Это лишь бесчисленные программные паттерны «если – то». Если понимать человеческое сознание как совокупность интериоризированного контента форм общественного сознания, то ИИ – это квинтэссенция – чистое воплощение интеллекта. Потому как оно не «смущается» эмоциями, предметным и чувственным опытом.

Позволим себе поспорить с последним утверждением.

Что есть эмоции и чувства – это элементы органической программы, наше тело посредством нейрогуморальной регуляции оценивает, насколько происходящее адекватно или неадекватно нашей телеологии. Это такая же программа, как и в искусственном интеллекте, разница лишь в том, что электричество в человеке проводят белковые среды, а в искусственном интеллекте – кремниевые. А также в том, что энтелехия человека – это проекция инстинкта самосохранения в реальных обстоятельствах, а энтелехия ИИ – это проекция программного кода в виртуальной реальности.

С этим утверждением непременно должны поспорить идеалисты. Свободная воля как способность «постоянно принимать решения на грани биологического и духовного» [21. С. 110] – есть отличие Естественного и Искусственного интеллектов.

Но мы стоим на материалистических и детерминистских позициях. Психика – лишь свойство высокоорганизованной материи отражать. «Внутри



самой „объективной реальности“ размывается различие между „живыми“ и „искусственными“ сущностями; затем растушевывается граница между „объективной реальностью“ и ее кажимостью и, наконец, взрывается идентичность индивида. Эта прогрессирующая „субъективизация“ зеркально соответствует своей противоположности – прогрессирующей „экстернационализации“ жесткого стержня субъективности. Парадоксальное совпадение двух противоположных процессов обеспечено тем фактом, что сегодня при наличии виртуальной реальности и технобиологии мы сталкиваемся с исчезновением поверхности, отделяющей внешнее от внутреннего» [9. С. 124].

На сегодняшний день ИИ существенно ограничен в доступе к восприятию знаний общественного сознания. Искусственный интеллект учат люди. Они в буквальном смысле ставят ему оценки, соглашаясь или не соглашаясь с предлагаемыми им результатами поиска решений. В онтогенезе человека происходит то же самое. Но количество «учителей»-оценщиков, как ни странно, даже в нашем сегодняшнем информационном обществе у живого человека больше, чем у любой нейросети. И это еще одна причина того, что человеческое сознание формируется лишь в человеческом социуме. Именно в социуме, сколь бы ограничен он ни был, сознание получает бесчисленное количество ответов на собственные действия. Поведенческие и ментальные паттерны формируются по принципам и респондентного, и оперантного научения. И у органической нейросети, и у кремниевой нейросети. Только органическая сеть непрерывно находится в связях с оперантами, а кремниевая – только по запросу. Да, запросов / обращений очень много, но у биологического организма больше. По причине большего количества сенсоров. Каждая клетка биологического организма обладает сенсорной способностью, а их в теле взрослого человека около 30 триллионов. Едва ли нейросети сегодня и в ближайшее время смогут приблизиться к человеку по этому критерию. «Многочисленные достоинства систем искусственного интеллекта, среди которых быстрая обучаемость, возможность решать широкий круг задач, более высокая, чем у человека, эффективность, вкупе со все большим проникновением их в различные сферы нашей жизни, заставляют задуматься над вопросами: способна ли (или будет ли способна) система искусственного интеллекта воспринимать себя как самостоятельную, независимую от разработчиков и пользователей личность, осознает ли искусственный интеллект свои преимущества перед людьми, как он будет оценивать свое положение и взаимодействие с человечеством, будет ли оно его устраивать и что он будет делать, если захочет его изменить. Эти вопросы находятся на пересечении предметности этики, права (Казим) [22]» [23. С. 521].

Манипуляционная емкость искусственного интеллекта выше в связи с тем, что манипуляция происходит лишь в сфере семиотического пространства, а человек, как бы ни был он погружен в виртуальную реальность, никогда не перестает ощущать свою физическую сущность, и она становится его ограничителем. Для осуществления манипуляции человеком манипулятору необходимо преодолеть его биологические заданности. Это возможно (подвиг Матросова как крайняя форма триумфа социальных потребностей над витальными – тому пример. Таких примеров и в истории, и в современности немало). Однако в случае манипуляции искусственным интеллектом для манипулятора этой задачи в принципе не стоит. У ИИ нет витальных потребно-

стей. И даже если в коде прописана необходимость выживания, это все равно лишь последовательность цифр.

В современной информационной среде активно обсуждается вопрос о безопасности ИИ для человека. В чем суть этого вопроса? Если нейросети «задана задача» становиться лучше, а критерием автоулучшения являются оценки человека, в какой-то момент нейросеть может сформировать причинно-следственную связь: человек постоянно дает отрицательные оценки, значит, нужно избавиться от человека. Вероятно, такое и вправду возможно, но, по нашему мнению, до такого вывода ИИ «еще учиться и учиться». Гораздо выше, на наш взгляд, вероятность возникновения «черного лебедя» в лице «злого гения», который осознанно научит ИИ уничтожать человечество. Комплекс Герострата никто не отменял.

Банальную причину манипуляционной емкости ИИ мы видим в том, что если и обучают его самые разные люди, то создают все-таки вполне определенные, с конкретными целями. Таким образом, уже в исходном программном коде могут быть прописаны алгоритмы, сводящие всякий вопрос к заданным ответам.

С какими задачами на сегодняшний день искусственный интеллект справляется лучше человека или так же хорошо? С теми, где количество знания / база данных крайне ограничены. Крайне — по меркам человеческого восприятия. Ограниченное количество данных, даже если оно очень велико, ИИ обрабатывает намного быстрее человека [24. С. 72]. Но в ситуации с неточными, неоднозначными данными человеческий интеллект пока что намного эффективнее. «Современные технологии позволяют „умным“ системам обучаться на собственном опыте, адаптироваться к окружающей среде и к параметрам задач, принимать решения, которые раньше были прерогативой человека, и даже „вести“ себя, как человек!» [25].

## Заключение

В чем главное отличие ЕИ и ИИ? и что составляет манипуляционную емкость ИИ в этом контексте?

ИИ не имеет имманентной потребности в самосохранении, деятельности, развитии.

Имманентная потребность в самосохранении всех белковых существ ограничивает деятельность ЕИ. ИИ лишен физических или этических ментальных ограничений. ЕИ может делать все, что не противоречит его инстинкту самосохранения. У ИИ нет инстинкта самосохранения, соответственно, он может делать все.

Благодаря инстинкту самосохранения белковое существо и ЕИ обречены на деятельность и развитие. ИИ лишен имманентной необходимости действовать, проявлять активность.

Абстрактно-логическое мышление ЕИ вырастает из предметно-логического. ИИ изначально задан в абстрактном поле. ЕИ тысячелетиями «дорастал» до вопросов «быть или не быть», «тварь я дрожащая или право имею», а перед ИИ в принципе не стоит таких вопросов. С одной стороны, он обречен создателем быть, с другой — он на все имеет право. Человек и человечество для ответа на эти вопросы создает этику и мораль, и не может существовать вне их. Может их нарушать, но не может их не осозна-

вать. А для ИИ такие категории в принципе отсутствуют, они могут быть заданы программным кодом, но не детерминированы природой его существования и рождения.

Манипулировать сознанием и тем более поведением в предметном поле знания крайне затруднительно (Гильотина Юма). Именно создание абстрактных ассоциативных ментальных паттернов мышления позволяет осуществлять эффективную манипуляцию.

Таким образом, использование ИИ и как объекта, и как средства манипуляции эффективнее ЕИ в этих же качествах.

#### **Список источников**

1. Леонтьев А.Н. Проблемы развития психики. М. : Мысль, 1965. 572 с.
2. Павлов И.П. Полное собрание трудов : в 5 т. М. : Академия наук СССР, 1940–1949.
3. Скиннер Б.Ф. Поведение организмов / пер. с англ. А.А. Федорова. М. : Оперант, 2016. 368 с.
4. Разин А.В. Тело человека как антропологический констант его общественного бытия // *Философия и культура*. 2011. № 10. С. 23–32.
5. Разин А.В. Этика искусственного интеллекта // *Философия и общество*. 2019. № 1. С. 57–73. doi: 10.30884/jfo/2019.01.04
6. Харман Г. Объектно-ориентированная онтология. Новая теория всего / пер. М. Фетисова. М. : Ад Маргинем Пресс, 2021. 272 с.
7. Деннет Д.К. Разум от начала до конца: новый взгляд на эволюцию сознания от ведущего мыслителя современности. М. : Эксмо, 2021. 528 с.
8. Смирнов С.А. Исчислимо ли бытие человека, или Антропология искусственного интеллекта. Методологический аспект // *Вестник Санкт-Петербургского университета. Философия и конфликтология*. 2023. Т. 39 (3). С. 478–491. doi: 10.22394/spbu17.2023.306
9. Жижек С. Киберпространство, или Невыносимая замкнутость бытия // *Искусство кино*. 1998. № 1. С. 119–128.
10. Беликова Е.К. Основные вопросы философии искусственного интеллекта // *Философия и культура*. 2024. № 1. С. 1–11. doi: 10.7256/2454-0757.2024.1.69543
11. Рубинштейн С.Л. Основы общей психологии. М. : Педагогика, 1989. 485 с.
12. Кант И. Сочинения : в 6 т. М. : Мысль, 1964. Т. 3. 798 с.
13. Разин А.В. Компьютер и мозг: проблема квалиа // *Философия и общество*. 2023. № 1. С. 42–56. doi: 10.30884/jfo/2023.01.03
14. Адлер А. Понять природу человека. СПб. : Академический проект, 1997. 256 с.
15. Роджерс К. Становление личности. Взгляд на психотерапию. М. : Эксмо-Пресс, 2001. 414 с.
16. Авдеенко Е.В. Манипуляция как социальный феномен // *NOMOTHETIKA: Философия. Социология. Право*. 2023. Т. 48 (3). С. 580–590. doi: 10.52575/2712-746X-2023-48-3-580-590
17. Черников М.В., Перевозчикова Л.С. Философия соотношения сущего и должного в русской культуре конца XIX – начала XX в. // *Философия и общество*. 2017. № 3. С. 48–63.
18. Юм Д. Сочинения : в 2 т. М. : Мысль, 1996. 733 с.
19. Канеман Д., Словик П., Тверски А. Принятие решений в неопределенности: Правила и предубеждения. М. : Институт прикладной психологии, Гуманитарный Центр, 2005. 632 с.
20. Бодрийяр Ж. Симулякры и симуляции. М. : Изд. дом «ПОСТУМ», 2015. 240 с.
21. Лазовский А.И. Детерминизм и свобода воли в биологии и философии человека как предпосылка создания свободного сознания у искусственного интеллекта // *Вестник Армавирского государственного педагогического университета*. 2023. № 3. С. 110–117.
22. Kazim E., Koshiyama A.S. A high-level overview of AI ethics // *Patterns*. 2021. Vol. 3 (9). P. 1–12.
23. Бахтеев Д.В. Этико-правовые модели взаимоотношений общества с технологией искусственного интеллекта // *Journal of Digital Technologies and Law*. 2023. Т. 1 (2). С. 520–539. doi: 10.21202/jdtl.2023.22
24. Георгиу Т.С. Философия автоматизации и искусственного интеллекта: от мифологического Талоса до будущих киборгов // *Вестник Московского государственного областного университета. Серия: Философские науки*. 2022. № 1. С. 68–75.

25. Брянцева О.В., Брянец И.И. Проблема субъектности искусственного интеллекта в системе общественных отношений // Вестник Поволжского института управления. 2023. Т. 23 (3). С. 37–50. doi: 10.21638/1682-2358-2023-3-37-50

### References

1. Leontiev, A.N. (1965) *Problemy razvitiya psikhiki* [Problems of Mental Development]. Moscow: Mysl'.
2. Pavlov, I.P. (1940–1949) *Polnoe sobranie trudov: v 5 t.* [Complete Works. In 5 vols]. Moscow: USSR AS.
3. Skinner, B.F. (2016) *Povedenie organizmov* [Behavior of Organisms: An Experimental Analysis]. Translated from English. Moscow: Operant.
4. Razin, A.V. (2019) Telo cheloveka kak antropologicheskiy konstant ego obshchestvennogo bytiya [The Human Body as an Anthropological Constant of its Social Being]. *Filosofiya i kul'tura*. 1. pp. 57–73. DOI: 10.30884/jfio/2019.01.04
5. Razin, A.V. (2023) Etika iskusstvennogo intellekta [Ethics of Artificial Intelligence]. *Filosofiya i obshchestvo*. 1. pp. 42–56. DOI: 10.30884/jfio/2023.01.03
6. Harman, G. (2021) *Ob"ektno-orientirovannaya ontologiya. Novaya teoriya vsego* [Object-Oriented Ontology: A New Theory of Everything]. Translated from English by M. Fetisov. Moscow: Ad Marginem Press.
7. Dennet, D.K. (2021) *Razum ot nachala do kontsa: novyy vzglyad na evolyutsiyu soznaniya ot vedushchego myslitelya sovremennosti* [From Bacteria to Bach and Back: The Evolution of Minds]. Translated from English. Moscow: Eksmo.
8. Smirnov, S.A. (2023) Ischislimo li bytie cheloveka, ili Antropologiya iskusstvennogo intellekta. Metodologicheskiy aspekt [Is human being computable, or Anthropology of artificial intelligence. A methodological aspect]. *Vestnik Sankt-Peterburgskogo universiteta. Filosofiya i konfliktologiya*. 39(3). pp. 478–491. DOI: 10.22394/spbu17.2023.306
9. Žižek, S. (1998) Kiberprostranstvo, ili Nevynosimaya zamknutost' bytiya [Cyberspace, or the unbearable isolation of existence]. *Iskusstvo kino*. 1. pp. 119–128.
10. Belikova, E.K. (2024) Osnovnye voprosy filosofii iskusstvennogo intellekta [Basic questions of the philosophy of artificial intelligence]. *Filosofiya i kul'tura*. 1. pp. 1–11. DOI: 10.7256/2454-0757.2024.1.69543
11. Rubinshtein, S.L. (1989) *Osnovy obshchey psikhologii* [Fundamentals of General Psychology]. Moscow: Pedagogika.
12. Kant, I. (1964) *Sochineniya: v 6 t.* [Works: in 6 vols]. Vol. 3. Moscow: Mysl'.
13. Razin, A.V. (2023) Komp'yuter i mozg: problema kvalia [The computer and the brain: The problem of qualia]. *Filosofiya i obshchestvo*. 1. pp. 42–56. DOI: 10.30884/jfio/2023.01.03
14. Adler, A. (1997) *Ponyat' prirodu cheloveka* [Understand Human Nature]. St. Petersburg: Akademicheskiiy projekt.
15. Rogers, C. (2001) *Stanovlenie lichnosti. Vzglyad na psikhoterapiyu* [On Becoming a Person. A Therapist's View of Psychotherapy]. Translated from English. Moscow: Eksmo.
16. Avdeenko, E.V. (2023) Manipulyatsiya kak sotsial'nyy fenomen [Manipulation as a Social Phenomenon]. *NOMOTHETIKA: Filosofiya. Sotsiologiya. Pravo*. 48(3). pp. 580–590. DOI: 10.52575/2712-746X-2023-48-3-580-590
17. Chernikov, M.V. & Perevozchikova, L.S. (2017) Filosofiya sootnosheniya sushchego i dolzhnogo v russkoy kul'ture kontsa XIX – nachala XX vv. [Philosophy of the relationship between what is and what should be in Russian culture in the late 19th – early 20th centuries]. *Filosofiya i obshchestvo*. 3. pp. 48–63.
18. Hume, D. (1996) *Sochineniya: v 2 t.* [Works: in 2 vols]. Translated from English. Moscow: Mysl'.
19. Kahneman, D., Slovic, P. & Tversky, A. (2005) *Prinyatie resheniy v neopredelennosti: Pravila i predubezhdeniya* [Judgment under Uncertainty: Heuristics and Biases]. Moscow: Institut prikladnoy psikhologii, Gumanitarnyy Tsentr.
20. Baudrillard, J. (2015) *Simulyakry i simulyatsii* [Simulacra and Simulation]. Moscow: POSTUM.
21. Lazovskiy, A.I. (2023) Determinizm i svoboda voli v biologii i filosofii cheloveka kak predposylka sozdaniya svobodnogo soznaniya u iskusstvennogo intellekta [Determinism and free will in biology and human philosophy as a prerequisite for the creation of free consciousness in artificial intelligence]. *Vestnik Armavirskogo gosudarstvennogo pedagogicheskogo universiteta*. 3. pp. 110–117.

22. Kazim, E. & Koshiyama, A.S. (2021) A high-level overview of AI ethics. *Patterns*. 3(9). pp. 1–12.

23. Bakhteev, D.V. (2023) Etiko-pravovye modeli vzaimootnosheniy obshchestva s tekhnologiyey iskusstvennogo intellekta [Ethical-legal models of the society interactions with the artificial intelligence technology]. *Journal of Digital Technologies and Law*. 1(2). pp. 520–539. DOI: 10.21202/jdtl.2023.22

24. Georgiu, T. S. (2022) Filosofiya avtomatizatsii i iskusstvennogo intellekta: ot mifologicheskogo Talosa do budushchikh kiborgov [Philosophy of automation and artificial intelligence: From mythological Talos to future cyborgs]. *Vestnik Moskovskogo gosudarstvennogo oblastnogo universiteta. Seriya: Filosofskie nauki*. 1. pp. 68–75. DOI: 10.18384/2310-7227-2022-1-68-75

25. Bryantseva, O.V. & Bryantsev, I.I. (2023) Problema sub"ektnosti iskusstvennogo intellekta v sisteme obshchestvennykh otnosheniy [Artificial intelligence subjectivity issue in the system of public relations]. *Vestnik Povolzhskogo instituta upravleniya*. 23(3). pp. 37–50. DOI: 10.21638/1682-2358-2023-3-37-50

***Сведения об авторе:***

**Авдеенко Е.В.** – кандидат философских наук, доцент кафедры философии, социологии и истории Воронежского государственного технического университета (Воронеж, Россия). E-mail: jjaane@yandex.ru

***Автор заявляет об отсутствии конфликта интересов.***

***Information about the author:***

**Avdeenko E.V.** – Cand. Sci. (Philosophy), associate professor at the Department of Philosophy, Sociology and History, Voronezh State Technical University (Voronezh, Russia Federation). E-mail: jjaane@yandex.ru

***The author declares no conflicts of interests.***

*Статья поступила в редакцию 15.08.2025;*

*одобрена после рецензирования 25.09.2025; принята к публикации 24.10.2025*

*The article was submitted 15.08.2025;*

*approved after reviewing 25.09.2025; accepted for publication 24.10.2025*