

Tokareva Anastasiya Aleksandrovna, Candidate of Cultural Studies, Deputy Head, Centre “All-Russian Patent Technical Library”, Federal Institute of Industrial Property. Postal address: 30, Berezhkovskaya Emb., G-59, GSP-3, Moscow, Russian Federation, 125993. Tel.: 8 (499) 240-64-13. E-mail: anastasiia.tokareva@rupto.ru

Evdokimova Maya Igorevna, Candidate of Philosophy, Head, Center for Interaction with Authorities, Federal Institute of Industrial Property, Associate Professor, Department of Intellectual Property, Griboyedov Moscow University. Postal address: 30, Berezhkovskaya Emb., G-59, GSP-3, Moscow, Russian Federation, 125993. Tel.: 8 (499) 240-33-55. E-mail: Maya-evd@mail.ru

DOI: 10.17805/zpu.2026.1.7

## Неклассическая машина: проблема контроля

С. В. ОБОЛКИНА

ИНСТИТУТ ФИЛОСОФИИ И ПРАВА УРО РАН

*В статье развивается взгляд на искусственный интеллект (ИИ) с позиции онтологии машины. Показано, что концепт машины в современной философии функционирует в качестве нового всеобщего, однако сама машина мыслится понятной по умолчанию. Такой подход не позволяет видеть специфику машины нового типа — генеративного ИИ. Внимание к машине как таковой и к новому типу машин позволяет более реалистично оценивать как перспективы, так и опасности нашего взаимодействия с ИИ. В первую очередь в работе поставлен вопрос: можно ли исключить фактор сделанности из представления об онтологии машины? На основе ключевых позиций теории машин и механизмов русской инженерной школы сделаны выводы, во-первых, о том, что фактор генезиса машины невозможно исключить из понимания сущности машины; во-вторых, что онтология машины в классическом ее понимании связана с минимизацией хаотических степеней свободы. Однако, обратившись к рассмотрению машины ИИ, мы должны подвергнуть сомнению это базовое для онтологии машины положение, потому что человечество построило новый вид машин. В нем фактор сделанности остается ключевым, но взаимоотношения с хаосом существенно меняются. Для анализа генеративного ИИ с точки зрения онтологии вводится концепт «неклассическая машина». Это машина, в которой хаос становится элементом ее механизмов. Раскрывается специфика неклассической машины. Рассмотрена парадигма «механической прозрачности», которая, как надеются цифровые инженеры, вернет машине ИИ характеристики контролируемости и онтологической сопряженности с когницией человека.*

*Ключевые слова: машина; неклассическая машина; онтология машины; генеративный ИИ; естественные абстракции; прозрачность ИИ*

### ВВЕДЕНИЕ:

#### МАШИНА КАК ФИЛОСОФСКАЯ ПРОБЛЕМА

Современный человек активно выстраивает сосуществование с ИИ. Для кого-то это повод для футурологического оптимизма; для кого-то актуальны неолуддистские настроения, инициированные прогнозами о «замене» человека этими системами. В данной работе мы сосредоточимся на другом ракурсе: ИИ — это в пер-

вую очередь машина. Основной тезис: современный тип ИИ связан с серьезной трансформацией онтологии машины, и внимание к новому типу машин позволяет более реалистично оценивать как перспективы, так и опасности сосуществования с ИИ.

Казалось бы, современную мысль нельзя упрекнуть в недостатке внимания к машине. Машина признается новым всеобщим (см., например, Николин, 2002). «Машинность» мыслится тем основанием, которое можно обнаружить за всяким видимым многообразием. В. Подорога предлагает онтологическое обобщение: «Теперь она (машина. — С. О.) осуществляет полный мимесис Реальности, давая нам реальное в машинном образе (сегодня — в цифровом. — С. О.)» (Подорога, 2010: 23). Л. Брайант формулирует принципы «панмеханизма» в качестве основной на сегодня, как он считает, философской стратегии: «Панмеханизм — это не просто утверждение, что бытие целиком состоит из машин, но что все взаимодействия являются механическими взаимодействиями. Конечно, мы еще не знаем, что это значит. Очевидно, что существует целая зоология машин, которую необходимо интенсивно исследовать» (Bryant, 2012). В футурологическом аспекте онтологический приоритет «машинной формы жизни» выступает одним из модусов трансгуманизма. В частности, в концепции Н. Ланда будущее видится в качестве такой реальности, которая «обнажилась», наконец, до своих машинных основ: «...наш человеческий камуфляж снимается, кожа с легкостью отрывается, обнажая сверкающую электронику. Информация струится в нас из Киберии — базы истинной революции, сокрытой в будущем от земной иммунополитики. Ровно в полночь века мы выйдем из наших логов, чтобы разобрать все системы безопасности, составляя завтрашний день» (Ланд, 2018: 34).

Во многом эти интенции питаются идеями шизоанализа Ж. Делёза и Ф. Гваттари, которые выстраивают свои положения относительно природы человека, социума и т. д. на основе метафизического нарратива о машине. Как показывает М. Хардт, их философия — в первую очередь именно онтология и лишь затем форма психо- и культуроанализа: «“Все есть машины, машины, соединенные с другими машинами”, это утверждение о природе реальности» (Hardt, 1993: 26). Делёз и Гваттари предлагают концепт «машины желания», в котором само бытование желания, его сущность имеет природу бессубъектного «машинного бессознательного»: «... машины желания — это действительно то же самое, что технические и социальные машины, но они являются, так сказать, их бессознательным: они проявляют и мобилизуют инвестиции желания, которые “соответствуют” сознательным или предсознательным инвестициям в интересы, политику и технологию определенной социальной сферы» (Deleuze, Guattari, 2009: 111). «Желание — это машина, и объект желания — это подсоединенная машина» (Делёз, Гваттари, 2007: 49–50). Поэтому «больше нет ни природы, ни человека, есть лишь процесс, который производит одно в другом и состыковывает машины» (там же: 14).

Таким образом, можно утверждать, что в современной онтологии активен тренд *машинизма*. Само понятие «машинизм» возникло в начале XX в.; одним из первых его использовал Ле Корбюзье в манифесте «Архитектура в эпоху машинизма» (Ле Корбюзье, 2017: 113). Гваттари использует его в качестве еще одного самоназвания шизоанализа — не только для того, чтобы указать на онтологическую роль концепта «машина», но и для того чтобы подчеркнуть его современные коннотации. В ответ на вопрос: «Чем новый машинизм, одним из основателей ко-

торого Вы являетесь, отличается от машинизма эпохи Просвещения, которое также было одержимо разными механизмами?» — Гваттари расширяет область значения понятия «машина»: «Машины, с которыми имеем дело мы, значительно более разнообразны: среди них, как я уже говорил, есть и технологические машины, но также машины лингвистические в том смысле, в каком этот термин понимает Хомский, урбанистические машины, мегамашины городов, есть также эстетические машины» (цит. по: Рыклин, 2002). Таким образом, машинизм — это мировоззренческая установка и философское направление, предлагающие понимание реальности за счет семантического расширения понятия машины за границы обозначения только технического артефакта. При этом стоит отличать современный тренд машинизма от нововременного механицизма — существует как общая концептуальная основа, так и принципиальная несводимость машинизма к механицизму. В данной работе, однако, мы оставляем эту тему без развития, отсылая к исследованию, посвященному этому вопросу (Оболкина, 2025).

Стоит отметить, что в таком расширительном понимании машины философия солидарна со значительной частью современной науки. Естествознание XIX–XXI вв. активно анализирует процессы трансформации энергии в терминах функционирования машины: говорится о процессах, в которых запасенная энергия/теплота реализуются в качестве работы; пища — это «топливо», а «запасенная энергия» высвобождается сожжением, окислением и т. п. И кибернетика в силу поставленных задач принципиально не проводит различия между машинной и немашинной реальностью. Н. Винер рассматривает машину как «устройство для преобразования входных сообщений в выходные» (Винер, 1966: 42). У. Росс Эшби, рассуждая о машинах, уточняет: «как построенные человеком, так и естественные» (Эшби, 1959: 49), т. е. по умолчанию предполагается, что существуют «естественные машины». Делается вывод о том, что «мы и только мы» решаем, что нам признавать за «похожее на машину» и что не признавать» (там же: 64–65).

Однако в качестве истока философской рефлексии о машине стоит все же задать вопросом: что позволяет нам быть уверенными в том, что *машина* — *всеобщее, а не особенное*? Обращает на себя внимание то, что именно элиминация фактора созданности, сделанности машины позволяет осуществляться онтологическим обобщениям машинизма. Не в последнюю очередь это связано с крайне богатым семантическим ресурсом самого слова «машина».

Изначально греческое слово «машина» (механэ) означало осадную машину, но очень быстро расширило свой семантический потенциал: «машина» стала обозначать человеческую «хитрость», «уловку», усиливающую его природу (Зарипов, 2015). «Машиной» могло называться нечто как аксиологически благое, так и нейтральное, а также пейоративное по своему восприятию: «механэ» часто использовалось в значении «козни» (что сохранилось в современном слове «махинация»). Таким образом, смысл «машины» изначально не замыкался в семантических границах обозначения технических устройств. Но все же только в современных текстах коннотации сделанности «уловки» исчезают из семантического спектра «машины». Если Л. Мамфорд посредством «машины» характеризует те черты социальных систем, которые связаны с категориями контроля, эффективности и власти в отношении государства (Мамфорд, 2001), т. е. так или иначе выходит на мотивы учреждения, то активно использующие его идеи Делэз и Гваттари понимают машину в контексте принципиально *бессознательной* и *бес-*

*субъектной* стихии производства желания (которая в итоге создает и машины государства тоже).

Пожалуй, единственным сектором анализа машины в теоретически-абстрактном ключе, не разрывающим при этом с фактором сделанности (созданности), выступают философская традиция механологии и научная дисциплина «Теория машин и механизмов» (ТММ).

В работе Ж. Лафита «Размышления о науке о машинах» (1932) (Lafitte, 1972) ставится вопрос о том, что это означает: понимать машину. Лафит говорит, что знания механики, физики, химии и инженерии, хотя и являются обширными и необходимыми для понимания машины, но не исчерпывают его. Исследователь предлагает для такого философского осмысления машины, которое не отрывалось бы от технического и естественно-научного ресурса понимания машины, но в то же время не сводилось бы к нему, название «механология». Этот подход и название использует Ж. Симондон, который настаивает на необходимости исследовать машину принципиально как артефакт. Фактор сделанности есть основание принципа непрерывности в нашем внимании к машине. Машина, по мысли философа, есть результат трансдуктивного процесса творчества, который приводит к технической индивидуации. Акт создания/изобретения машины есть момент единения гетерогенных сред природы, человека, техники (Симондон, 2012).

В таком понимании машина как предмет философского исследования отличается от предмета философии техники. Философское исследование природы техники — это прежде всего исследование одной из форм мироотношения, познания и деятельности человека. Философ техники П. К. Энгельмейер пишет по поводу употребления слова «техника» в ракурсе этой дисциплины: «...будем употреблять в самом широком смысле, а именно означаем им все человеческие знания, направленные на практические цели, также все умения» (Энгельмейер, 1887). Технэ еще со времен античной эпохи мыслится как одна из форм проявленности истины *для человека*. Механология же уточняет, что предметом ее философского исследования выступает не техника, а именно технический объект. В данном исследовании мы предлагаем развить этот ракурс внимания к машине, углубляя его важнейшими положениям ТММ.

Для ТММ машина — это система механизмов: «Механизмом называется система материальных тел, предназначенных для преобразования движения одного или нескольких тел в требуемые движения остальных» (Недоводеев, 2012). Механизм рассматривается как форма и способ такого попарного соединения элементов, при котором они ограничивали бы движение друг друга согласно конкретной инженерной задаче. Ключевым примером и образом базового механизма в ТММ выступает зубчатая передача.

Элементы механизма — детали (изделие, изготовленное без сборочных операций). Детали, соединенные между собой неподвижно или с помощью упругих связей, образуют отдельное звено. Два звена, соединенных друг с другом подвижно, образуют кинематическую пару. Именно в отношении кинематической пары развиваются такие положения ТММ, которые, с одной стороны, выводят нас на максимально абстрактное понимание машины, с другой, позволяют не терять из виду фактор созданности и конкретику инженерной практики. Именно к этому срезу знания о машинах мы обратимся далее, ставя перед собой задачу философского исследования феномена машины для того, чтобы постараться понять: не имеем ли

мы в случае с современным ИИ преодоление границ самого феномена машины? Интуитивно мы чувствуем, что ИИ — «не просто машина», но что фактически стоит за этой интуицией?

#### МАШИНА И ВОПРОС О СТЕПЕНЯХ СВОБОДЫ

Русская и советская инженерная школа утвердили в науке о машинах теоретический подход, связанный с анализом степеней свободы. Во второй половине XIX в. Ф. Грасгоф одним из первых пользуется понятием степеней свободы для определения кинематических пар в качестве базового элемента всякого механизма (Боголюбов, 1976: 317). Однако эта идея еще не была в достаточной степени абстрактной, не имела математического характера. Специфика подхода российских исследователей во многом была связана с преодолением сугубо описательного подхода к анализу машин и механизмов. Исследования П. А. Чебышева, О. И. Сомова и их учеников (Петербургская школа математики и механики) заложили основу для формирования не только теоретического анализа, но и синтеза машин и механизмов. Сын О. И. Сомова П. О. Сомов делает категорию «степень свободы» принципиальным основанием всякого исследования структуры механизмов (начало положено в работе «О степенях свободы кинематической цепи», 1887). И, как отмечает И. И. Артоболевский, вклад П. О. Сомова, а именно критерий подсчета степеней свободы для анализа и синтеза механизмов, является базой русской и советской школы ТММ (Артоболевский, 1943). Позднее профессор ТТИ<sup>1</sup> А. П. Малышев развил идеи Сомова, и в итоге формула Сомова — Малышева вошла в качестве обязательного элемента в ТММ<sup>2</sup>.

Малышев берет на вооружение определение Сомова: «...механизмом называется такая кинематическая цепь, в которой все точки описывают определенные траектории, т. е. каждое звено которой не имеет больше одной степени свободы» (Малышев, 1923: 10). В самом широком смысле «степени свободы» — это понятие, анализирующее независимые направления развития системы и независимые переменные, характеризующие состояние системы. Число степеней свободы равно минимальному количеству переменных, которое необходимо для полного описания состояния системы. Таким образом, речь идет о *возможностях контролировать хаос реального*.

В обобщенно-математическом смысле хаос — это невозможность вывести закономерности, которые позволили бы прогнозировать каждое последующее изменение траектории частицы по предыдущему ее состоянию. Минимум хаоса (хаотических степеней свободы) предполагает развитие системы по конкретной траектории; максимум хаоса означает, что система удаляется от равновесия по непредсказуемому сценарию. Основа выстраивания машины — это точное знание всех параметров движения во всех кинематических парах и их ансамблях. Минимум знания, исчерпывающе описывающего это движение (на абстрактном уровне это означает знание о координатах точек) = Минимум хаоса. Эти моменты и составляют содержание фактора выстроенности, сделанности машины.

Стоит подчеркнуть онтологическое обобщение, которое делает Малышев в своей работе: исследователь говорит о том, что всякая движущаяся точка всегда имеет лишь одну степень свободы: «В природе фактически движение происходит только с одной степенью свободы. <...> Это относится ко всем без исключения телам нашего мира, а потому нужно вообще заключить, что в нашем прост-

ранстве есть движение только с одной степенью свободы» (там же: 9). Это обобщение может показаться схожим с онтологическими позициями машинизма, поскольку не дает нам теоретических оснований различения машинной и немашинной реальности. Однако, будучи механиком, Малышев не может «путать» ситуации *анте* и *постфактум*: выстраивая механизм, мы *заранее* знаем эту одну степень свободы, тогда как в отношении просто реальности знаем *постфактум*, когда движение уже осуществилось в силу выбранных системой обстоятельств. Поэтому постулируя, что «движение вообще не имеет больше одной степени свободы», Малышев подчеркивает: «...в определении Сомова нужно подразумевать тот смысл, что получаемая в механизме одна степень свободы обеспечивается только теми условиями связи, которые нам известны и которыми мы распоряжаемся. Если этого не будет, то не будет и механизма, хотя бы в цепи была только одна степень свободы под влиянием каких-нибудь условий связи, назначенных природой» (там же: 10)

Необходимость создавать механизм (сколь бы ни было сложный) из звеньев с фиксированными степенями свободы и есть основание машины. Машина — в первую очередь пространство нашего контроля, и лишь в силу этого момента она может быть пространством производства, циклов работы и характеризоваться в качестве потоков и срезов и т. п.

Но далее, обратившись к рассмотрению машины ИИ, мы должны подвергнуть сомнению это базовое для онтологии машины условие: ограничение степеней свободы вплоть до минимума хаоса. Потому что человечество построило новый вид машин. В нем фактор сделанности остается ключевым, однако взаимоотношения с хаосом существенно меняются.

#### ИИ КАК НЕКЛАССИЧЕСКАЯ МАШИНА

Несмотря на то что вышесказанное касалось механических машин, мысль о степенях свободы применима и в отношении компьютерных технологий. И речь не только об аппаратном уровне (хард), но и о программной части этой машины (софт).

Важность программной инженерии очень пронизательно отмечала уже Ада Лавлейс, которая считается первым программистом (см., например, Боголюбов, 1976: 186). Единственная научная работа А. Лавлейс является примечаниями к статье Л. Ф. Менабреа, описывающей аналитическую машину Ч. Беббиджа (Menabrea, 1842). В одном из Примечаний Лавлейс отмечает: «Изучая работу аналитической машины, мы обнаруживаем, что особая и независимая природа рассуждений, которые во всем математическом анализе относятся к *операциям*, в отличие от *объектов*, над которыми производятся операции, и от *результатов* операций, выполняемых над этими объектами, очень четко определена и обособлена» (Примечание А). А сам Менабреа подчеркивает: «Хотя машина сама по себе не является мыслящим существом, ее можно рассматривать как существо, реализующее представления разума. Карточки получают отпечаток этих концепций и передают различным механизмам, составляющим машину, команды, необходимые для их работы (там же).

«Представления разума», оставляющие «отпечаток» на механизмах, неизбежно должны быть онтологически подобны кинематическим парам: в них тоже должен реализовываться принцип минимизации степеней свободы. Но «представле-

ния человеческого разума» и естественный язык не обладают данным качеством сами по себе. Лишь исключительно искусственным образом и локально человек может преобразовывать эту стихию в класс формализованных языков. Благодаря логике и математике он реализует то самое подобие с «кинематическими парами», а именно — с переключениями транзисторов.

Трудно не восхищаться масштабом выстроенных систем «солидарности» языкового и аппаратного, когда речь уже идет о машине под названием интернет. Это многоуровневая архитектура<sup>3</sup>, в которой есть «русла» для потоков информации и есть их наполнение (WWW). Вся эта сложная система алгоритмов на базовом уровне остается той самой классической машиной, которая выстраивается благодаря снижению степеней свободы, путем минимизации хаоса. Но в ходе развития компьютерных технологий появилась машина, для которой *хаос оказывается элементом механизма*. Ее вполне логично назвать неклассической, особенно если вспомнить о развитии неклассической логики, неклассических тенденций в философии, неклассических и постнеклассических новаций в естествознании, связанных с актуализацией вероятностного детерминизма, анализом нелинейных процессов, теорией хаоса и т. д. Речь и системах ИИ, основанных на технологиях нейронной сети.

Стоит подчеркнуть, что инженерия ИИ изначально не стремилась разорвать с парадигмой классической машины. Еще относительно недавно преобладали разработки ИИ, основанного на правилах: так называемый Good old fashioned artificial intelligence (GOF AI) — Старый добрый (или старомодный) искусственный интеллект. Суть этих классических разработок ИИ заключалась в попытках скопировать в машинных процессах основные компетенции человеческого разума: грамматические структуры языка, формирование суждений, распознавание образов. Неслучайно широко распространенное определение ИИ до сих пор несет следы идеи уподобления, копирования или сходства с интеллектом человека. В частности, в Национальной стратегии развития искусственного интеллекта на период до 2030 года (в ред. Указа Президента РФ от 15.02.2024 №124) ИИ был определен как комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая поиск решений без заранее заданного алгоритма) и получать при выполнении конкретных задач результаты, сопоставимые с результатами интеллектуальной деятельности человека или превосходящие их.

Однако попытки зафиксировать свойства человеческого интеллекта в качестве строгих формализмов требовали настолько масштабных и сложных построений цифровой инженерии, что довольно быстро стали видны проблемы с пределами вычислимости. Отчетливо обозначила себя и перспектива экспоненциального роста ошибок. Поэтому на основе когнитивных исследований сформировалась парадигма коннекционизма (от лат. «связь»): сначала идея о связи (сети) нейронов выступила принципом анализа работы мозга, и параллельно она стала ведущей концепцией в инжиниринге ИИ. При этом стратегия архитектуры нейронной сети неизбежно двигалась в направлении вероятностного подхода (начиная с концепции перцептрона Ф. Розенблатта): инженерные решения в сфере ИИ отошли от выстраивания формализмов, кодирующих каждый вычислительный акт. Работа с вероятностными алгоритмами и методами нечеткого управления легла в основу современного варианта ИИ: машины, способной на основе базового обучения к са-

мообучению и к демонстрированию результатов, обладающих определенной новизной и непредсказуемостью.

После этапа обучения нейросетям приходится иметь дело с данными как «диким множеством». Собственно, только ради этого они и создаются: к примеру, ИИ-помощники транспортных систем имеют дело с данными, которые фиксируют не контекст-модель, но контекст реального движения; чат-боты (LLM, т. е. большие языковые модели) обрабатывают запросы пользователей и данные из интернета. Эти хаотические, т. е. специально не подготовленные, данные встраиваются в распределение весов и влияют на дальнейшую работу ИИ. И чем «глубже» (технический термин, характеризующий количественную характеристику слоев нейросети) — а значит, «умнее» — нейросеть, тем сильнее проявляется в ее механизмах влияние фактора «дикого множества». Неструктурированные (хаотические) данные не просто обрабатываются детерминированным цифровым механизмом, но и выступают частью этого механизма. Обобщая это положение, можно сказать, что в какой-то степени и строго определенным образом, но хаос является элементом механизмов этого типа машин.

Для сравнения классической и неклассической машины можно привести ситуацию взаимодействия механизма мельницы и зерна. Зерно (сырье как хаос, «дикое множество») обрабатывается механизмом, что так или иначе влияет на кинетические и динамические характеристики элементов механизма за счет трения и т. п. Но зерно не встраивается непосредственно в структуру самого механизма, тогда как в нейросетях данные непосредственно участвуют в работе цифрового механизма. Для мельницы ситуация, когда зерно становилось бы частью механизма мельницы, — это абсурд или фантастика, для нейросети — норма. Поэтому имеет смысл обозначить дистинкцию классической и неклассической машины по отношению работы с хаосом. Понять характер сопряжения с хаосом в неклассической машине важно, поскольку от этого зависит и характер нашего взаимодействия с ИИ.

Первый момент участия хаоса в механизмах ИИ связан с тем, что LLM, т. е. программы, распознающие естественный язык и генерирующие ответы на таком языке, встраиваются в опоре на статистические методы учета частоты (частотностей) появления слов в предложении, тексте и т. д. Теория вероятностей по определению связана с категориями «случайность» и «хаос». Однако Б. Мандельброт предлагает использовать в отношении статистически интерпретируемой неоднородности данных концепты «прирученная случайность», «одомашненный хаос». Мыслитель долгое время исследовал шумы, турбулентность, биржевые процессы и в итоге настаивает, что у случая есть несколько различных состояний, и случайное, взятое в большом масштабе, всегда создает некоторую неслучайную, или «монотонную», закономерность. Только поэтому хаос может стать объектом научного внимания и математических манипуляций. В то же время «стихийный случай» не имеет к теории вероятностей прямого отношения. По словам Б. Мандельброта, «...“вероятности”, в зависимости от обстоятельств, либо преувеличивали, либо преуменьшали указанную “стихийность”» (Мандельброт, 2004: 90), и основной способ «одомашнивания» хаоса — смена оптики. «Если рассматривать различные выборки такого шума с достаточно близкого расстояния, то кажется, что они очень отличаются друг от друга — как и должно быть по самому определению случая. И наоборот, издали они очень похожи друг на друга и кажутся очень “монотон-

ными» (там же: 104). С таким «монотонным» воплощением хаоса имеет дело нейросеть, обученная ориентироваться (благодаря системе весов и базовому обучению) на «видимое издали», в больших масштабах. Нейросеть реализует «усреднение» данных. Будучи обученной на статистически значимых данных, она обрабатывает хаос реального в той мере, в какой он может быть обработан вероятностными алгоритмами, т. е. стать «монотонным».

На аппаратном уровне в механизм новой машины по-прежнему не может быть встроен хаос — мощные защиты от подобного воздействия составляют важную часть инженерного искусства. Но мы имеем дело с машиной, часть механизмов которой имеет идеальную природу, и конкретные математические методы «одомашнивают» хаос, т. е. могут позволять «абсорбировать» без сбоев системы «истинную случайность», не отменяя «дикое множество».

Но есть еще один уровень (модус) хаоса, выступающий элементом механизма неклассической машины, который является куда более сложной проблемой и для осмысления, и для практических решений, и в итоге для нашего сосуществования с ИИ. Этот модус хаоса также закономерно связан с «идеальной» частью механизмов неклассической машины: речь о базовых особенностях той онтологии, на которой выстраивается программная инженерия.

Наш образ реальности дан в языке. Основным инструментом выстраивания программного обеспечения выступает тот же язык, но «дистиллированный» до строгих логических конструкций. Казалось бы, в нем нет места хаосу. Но лазейка для него существует в качестве проблемы сопряжения языка и реальности.

Античная философия сделала ставку на их изоморфизм, который лежит в основе человеческой когниции<sup>4</sup>. Аристотель (говоря расширительно, т. е. подразумевая и платоновский дискурс, и многочисленных последователей философа) выстроил очень устойчивую онтокогнитивную позицию: мышление отражает реальность и на значительную свою часть состоит из абстракций. Но уже средневековые мыслители подвергли серьезной ревизии позиции аристотелевской онтологии. Проблемы универсалий и двойной сигнификации снова и снова ставили под вопрос позицию онтолингвистического изоморфизма. Встал вопрос: что является единицей существующего — имя или положение дел, фиксируемое суждением? Решения предлагаются и сегодня. Например, Б. Рассел показывает, что индивид, вещь — это повторяющийся в нашем восприятии комплекс качеств, который фиксируется языковой конструкцией. Доязыковое знание реальности дано в жесте остенсии, реализуется в указательном местоимении, и это, пожалуй, наиболее близкий контакт с реальностью. А. Коржибски, критикуя аристотелевскую систему когниции, показывает, что абстракции активны в первую очередь в языке, и потому в нас глубоко укоренилась потребность видеть за словами вещи (Моисеев, Коржибски, 2025). Мы, утверждает мыслитель, фиксируем индивидуальные объекты в некоей постулируемой идеальной изоляции, причем существительными-абстракциями выступает то, что, онтологически говоря, существительными не является: например, «тепло», «цвет» и т. п. То есть субъект-предикатная логика не соответствует структуре мира и нашего аффективного (т. е. несемантического) контакта с ним, это своего рода лингвистическая привычка.

В нашу задачу не входит глубокое погружение в проблемы аристотелевской метафизики и ее преодоления, мы лишь подчеркнем: человек (преимущественно западной культуры) с помощью логики, которая является плоть от плоти продуктом

этой метафизики, инсталлировал ИИ именно ту форму когниции, которая опирается на пусть и неявную, но убежденность в онтолингвистическом изоморфизме и в которой преобладают абстракции. Поэтому онтология ИИ отягощена трудностями, которые снова и снова пытается разрешить и философия, и цифровая инженерия.

Машина, таким образом, «мыслит» абстракциями. Это касается и классической, и неклассической машины, но с важными отличиями.

В классической машине (например, в логической машине, в классическом программном обеспечении, в ИИ, построенном на правилах, т. е. во всех цифровых машинах до самообучающегося ИИ) все акты ее механизмов прописаны в коде. Соответственно, набор абстракций этой машины задан самим программным обеспечением. В нем нет места хаосу, т. е. всегда есть возможность вывести закономерности, которые позволяют прогнозировать каждое последующее изменение работы системы.

Однако неклассическая машина — нейросеть (например, LLM) — принципиально создана так, чтобы после базового обучения абстракциям она должна формировать их сама из массива данных (т. е. «дикого множества»). И часто эти абстракции отличаются от тех, которые использует человек в качестве «опорных точек» восприятия реальности. Машине вообще нет нужды в выстраивании картины реальности, для нее важен лишь фактор оптимизации обработки данных. Какие-то измеримые свойства объектов (данных, которые обрабатываются машиной) концептуализируются машиной в качестве абстракций. Они являются результатом сведения множества характеристик к каким-то признакам (поэтому абстракции нейросетей называются обычно «признаками»). Исследователи подчеркивают, что «features (признаки) — это фундаментальные единицы, с помощью которых нейросети кодируют знания, основные компоненты представления НС (нейросетей. — С. О.)» (Bereska, Gavves, 2024). В идеале признаки должны были бы назначаться (соответствовать) отдельным нейронам, но в силу экономии ресурсов это практически неосуществимо. В итоге один нейрон оказывается связан с несколькими абстракциями, и возникает специфичная полисемантическая нейронная сеть: «Нейронные сети представляют больше признаков, чем у них нейронов, за счет кодирования признаков в перекрывающихся комбинациях нейронов» (там же). Это обстоятельство имеет онтологические последствия: нейросети воспроизводят процессы, схожие с формированием естественных абстракций у человека, однако абстракции, созданные машинами, иногда весьма существенно отличаются от тех, которые формирует наша когниция. Например, нейрон может активизироваться, реагируя как на кошку, так и на автомобиль, — для него это будет в какой-то момент одной и той же абстракцией. В цифровой инженерии это называется «неортогональность признаков».

Следует понимать, что для самой машины ИИ этот момент не выступает каким-то диссонансом. Она лишь оптимизирует данные. Но для нас это оказывается увеличением степеней свободы машины, непредсказуемостью ее поведения. Можно сказать, что нейросети все-таки создают *свои модели реальности* в качестве побочного продукта циклических процессов самообучения. В первую очередь это касается LLM, которые «...формируют сложные корреляции на основе данных наблюдений, но не обладают способностью разрабатывать причинно-следственные модели мира из-за отсутствия доступа к эмпирическим данным» (Pearl, 2009). При

этом машины ИИ все больше погружаются в среду реального мира, влияя на него обратной динамикой посредством контекстного обучения и в ходе контактов с пользователями.

Инженеры закладывают основания подобных онтологических «вольностей» уже тем, что делают ставку на теорию естественной абстракции. Но инженеры бьют тревогу, поскольку перспектива онтологической «непрозрачности» означает минимизацию контроля (Toward Transparent AI ... , 2023). Анализ только исходной модели не позволяет инженерам гарантировать надежную интерпретируемость представлений нейросети. Выражаясь упрощенно, мы не знаем, что нейросеть решит усовершенствовать, на каком этапе станет оптимизировать данные и выдавать результат согласно своему «образу реальности». Поэтому сегодня решается проблема того, как, не превращая неклассическую машину в классическую, сделать ее все же более машиной, т. е. реализовать ее работу в качестве нашей зоны контроля. Эта проблема является одним из важных моментов так называемой проблемы «черного ящика» (Black Box Problem), хотя и не сводится к ней (см., например, Zednik, 2021).

Инженеры отмечают, что основная проблема заключается в том, что нейросети, скорее, выращиваются, чем выстраиваются. А отслеживание их функциональности превращается в поведенческий анализ, а не редукционистское исследование детерминированных процессов. «Разрыв между онтологиями человека и модели может увеличиваться по мере развития архитектуры, повышая ее непрозрачность», поэтому «получение механистического представления о внутренней работе систем ИИ представляется крайне важным для обеспечения безопасности ИИ по мере разработки более мощных моделей» (Bereska, Gavves, 2024). Какие-то инженерные решения в рамках парадигмы «механической прозрачности» делают ставку на увеличение количества моносемантических нейронов, другие направлены на внедрение алгоритмов контроля путем выделения наложенных друг на друга признаков; предлагаются инструменты для интерпретации того, как системы ИИ принимают решения.

С позиции внешнего наблюдателя за этими процессами минимизации хаоса в ситуации его обязательности для неклассической машины можно попытаться сделать прогноз: будущее все-таки за специализированными системами ИИ, а не большими языковыми моделями. Если последние уже сегодня галлюцинируют, то специализированные модели, обученные на решение конкретных задач (например, в медицине, агротехнологиях и т. д.), хотя и требуют интерпретации естественного языка, не нуждаются в формировании абстракций разного уровня, чтобы синтезировать нечто похожее на когницию человека (хотя они очень чувствительны к качеству специализированных данных, но это уже вопрос, выходящий за пределы поставленных здесь проблем).

#### ЗАКЛЮЧЕНИЕ

Наше сосуществование с ИИ должно быть фундировано пониманием ИИ в первую очередь как машины, а не некой сущности, которая «умеет думать». В «лице» современного ИИ мы столкнулись не с угрозой сознания, превышающего наше. Мы изобрели *неклассическую машину*, в которой хаос выступает необходимым элементом ее механизмов. Данный концепт является важным для того, чтобы в ряду машин выделить ту, которая перестает быть «зоной контроля». Неклассическая

машина радикализирует онтологические «двусмысленности» нашего языка, что является вызовом не только для инженеров, но и для философов. «Склонность» к абстракциям, воплощенная в качестве цифровых механизмов, провоцирует формирование «образов реальности» вне и помимо человеческих интуиций. Поэтому философское исследование ИИ должно развивать понимание машины, модусов ее онтологии. Что оказывается одновременно исследованием онтологических ресурсов нашей когниции.

#### ПРИМЕЧАНИЯ

<sup>1</sup> Томский технологический институт Императора Николая II, сейчас Томский политехнический университет (ТПУ).

<sup>2</sup> В том числе в учебные курсы (см., например, Белов, Сорокин, 2020: 38–41).

<sup>3</sup> Многоуровневая архитектура — набор правил, который описывает процесс взаимодействия устройств в Сети. Open Systems Interconnection model (OSI model) выступает первой стандартной моделью в области сетевых коммуникаций.

<sup>4</sup> Когниция — «это любой процесс (сознательный или неосознанный), связанный с получением информации, знаний, их преобразованием, запоминанием, извлечением из памяти, использованием. Это и восприятие мира, и наблюдение, и категоризация, и мышление, и речь, и воображение и многие другие психические процессы или их совокупность» (Кубрякова, 2001: 39–49). В данном случае это понятие важно своим максимально широким пониманием познавательных процессов.

#### СПИСОК ЛИТЕРАТУРЫ

Артоболовский, И. И. (1943) Русская школа по теории механизмов и машин // Известия АН СССР. Вып. 7. С. 3–14.

Белов, М. И., Сорокин, С. В. (2020) Теория механизмов и машин : учеб. пособие. 2-е изд. М. : РИОР: ИНФРА-М, 2020. 322 с.

Боголюбов, А. Н. (1976) Теория механизмов и машин в историческом развитии ее идей. М. : Наука. 466 с.

Винер, Н. (1966) Творец и робот. Обсуждение некоторых проблем, в которых кибернетика сталкивается с религией. М. : Прогресс. 104 с.

Делёз, Ж., Гваттари, Ф. (2007) Анти-Эдип: Капитализм и шизофрения. Екатеринбург : У-Фактория. 672 с.

Зарипов, Я. В. (2015) Машина: анализ семантики в диахронической перспективе [Электронный ресурс] // Материалы Всероссийской школы молодых византинистов NEANIA. URL: <https://www.hse.ru/data/2015/10/03/1077230538/Зарипов%20Я.В.pdf> (дата обращения: 03.08.2023).

Кубрякова, Е. С. (2001) О когнитивной лингвистике и семантике термина // Вестник ВГУ. Серия: Лингвистика и межкультурная коммуникация. №1. С. 4–10.

Ланд, Н. (2018) Киберготика // Соч. В 6 т. Т. 1. Пермь : HylePress. 208 с.

Ле Корбюзье (2017) Новый дух в архитектуре. М. : Strelka Press. 113 с.

Мальшев, А. П. (1923) Анализ и синтез механизмов с точки зрения их структуры [Электронный ресурс] // Известия Томского технологического института [Известия ТТИ]. Т. 44. URL: <http://earchive.tpu.ru/handle/11683/7368> (дата обращения: 03.08.2023).

Мамфорд, Л. (2001) Миф машины. Техника и развитие человечества. М. : Логос. 408 с.

Мандельброт, Б. (2004) Фракталы, случай и финансы. М. — Ижевск : НИЦ «Регулярная и хаотическая динамика». 256 с.

Моисеев, Д. В., Коржибски, А. (2025) Наука и здравомыслие — 2025. Обновление методологии общей семантики в XXI веке. Издательские решения [Электронный ресурс]. URL: [https://books.yandex.ru/books/sDFrDlzi/read-online?with\\_payment\\_widget=true](https://books.yandex.ru/books/sDFrDlzi/read-online?with_payment_widget=true) (дата обращения: 03.08.2025).

Недоводеев, В. Я. 2012. Курс лекций по теории механизмов и машин для машиностроительных специальностей. Ульяновск : УлГТУ. 83 с

Николин, В. В. (2022) Машинность как всеобщий принцип воспроизводства: Онтологический аспект экспансии машины в немашинные сферы бытия : автореф. дис. ... д-ра филос. наук. Омск : ОмГПУ. 32 с.

Оболкина, С. В. (2025) Краткий очерк машинизма // Идеи и идеалы. №3–1. С. 124–143.

Подорога, В. (2010) Homo ex machina. Авангард и его машины. Эстетика новой формы // Логос. №1. С. 22–50.

Рыклин, М. (2002) Машины желания и просто машины [Электронный ресурс] // Деконструкция и деструкция: Беседы с философами. URL: <https://schizopromenade.blogspot.com/2014/07/guattari.html> (дата обращения: 03.08.2025).

Симондон, Ж. (2012) О способе существования технических объектов [Электронный ресурс] // Транслит: литературно-теоретический журнал. №9. URL: <http://www.trans-lit.info/materialy/9-vypuski/zhilber-simondon-o-sposobe-sushhestvovaniya-tehnicheskikh-obektov> (дата обращения: 03.08.2025).

Степанов, М. (2010) Аппаратное понимание медиа // Вестник ЛГУ им. А. С. Пушкина. №3. С. 194–199.

Энгельмейер, П. К. (1887) Экономическое значение современной техники. М. : Русская типография. 51 с. [Электронный ресурс] URL: [https://archive.org/details/ekonomicheskoe\\_znachenie\\_sovremennoy\\_tekhniki/page/2/mode/2up](https://archive.org/details/ekonomicheskoe_znachenie_sovremennoy_tekhniki/page/2/mode/2up) (дата обращения: 03.08.2025).

Эшби, У. Р. (1959) Введение в кибернетику. М. : Издательство иностранной литературы. 432 с.

Hardt, M. (1993) Gilles Deleuze: an Apprenticeship in Philosophy. Minneapolis : University of Minnesota Press. 457 p.

Bereska, L., Gavves, E. (2024) Mechanistic Interpretability for AI Safety — A Review // Transactions on Machine Learning Research. URL: <https://leonardbereska.github.io/blog/2024/mechinterpreview/> (дата обращения: 03.08.2025).

Bryant, L. (2012) Machine oriented ontology. URL: <https://larvalsubjects.wordpress.com/2012/06/21/machine-oriented-ontology-towards-a-pan-mechanism/> (дата обращения: 03.08.2025).

Deleuze, G., Guattari, F. (2009) Balance Sheet for Desiring Machines // Guattari, F., Lotringer S. Chaosology: texts and interviews, 1972–1977. Los Angeles, CA : Semiotext(e). P. 90–115.

Lafitte, J. (1972) Reflections on the Science of Machines. London : Mechanology Press.

Menabrea, L. F. (1842) Invented by Charles Babbage. With notes upon the Memoir by the Translator Ada Augusta, Countess of Lovelace from the Bibliothique Universelle de Genive, October, №82. URL: <https://fourmilab.ch/babbage/sketch.html> (дата обращения 03.08.2025).

Pearl, J. (2009) Causality. Cambridge University Press.

Toward Transparent AI: A Survey on Interpreting the Inner Structures of Deep Neural Networks (2023) / T. Ruker, A. Ho, S. Casper and D. Hadfield-Menell. Raleigh, 8–10 February. P. 464–483. URL: <https://doi.org/10.1109/satml54575.2023.00039> (дата обращения: 03.08.2025).

Zednik, C. (2021) Solving the Black Box Problem: A Normative Framework for Explainable Artificial Intelligence // Philosophy & Technology. №34. P. 265–288.

*Дата поступления: 10.01.2026 г.*

THE NON-CLASSICAL MACHINE:  
THE PROBLEM OF CONTROL

S. V. OBOLKINA

RAS URAL BRANCH INSTITUTE OF PHILOSOPHY AND LAW

The article develops a view of artificial intelligence (AI) from the perspective of machine ontology. It demonstrates that the concept of a machine in contemporary philosophy functions as

a new universal, yet the machine itself is conceived as self-evident. This approach obscures the specificity of a new type of machine — generative AI. Focusing on the machine as such and this new type of machine allows for a more realistic assessment of both the prospects and the dangers of our interaction with AI. First and foremost, the paper asks: can the factor of manufacturedness be excluded from the concept of machine ontology? Drawing on key tenets of the theory of machines and mechanisms from the Russian engineering school, the author concludes, firstly, that the factor of machine genesis cannot be excluded from understanding the essence of a machine; and secondly, that machine ontology, in its classical understanding, is associated with the minimization of chaotic degrees of freedom. However, when considering the AI machine, we must question this fundamental assumption of machine ontology, for humanity has built a new type of machine. In this new type, the factor of manufacturedness remains key, but its relationship with chaos has changed significantly. To analyze generative AI from an ontological perspective, the concept of a «non-classical machine» is introduced. This is a machine in which chaos becomes an integral part of its mechanisms. The specific characteristics of a non-classical machine are explored. The paradigm of «mechanical transparency» is considered, which, digital engineers hope, will restore controllability and ontological congruence with human cognition to AI machines.

Keywords: machine; non-classical machine; machine ontology; generative AI; natural abstractions; AI transparency

#### REFERENCES

Artobolevsky, I. I. (1943) Russkaya shkola po teorii mekhanizmov i mashin. *Izvestiya AN SSSR, OTN, iss. 7. pp. 3–14.*

Ashby, W. R. (1959) *Vvedeniye v kibernetiku*. Moscow, Publishing House of Foreign Literature. 432 c.

Belov, M. I., Sorokin, S. V. (2020) *Teoriya mekhanizmov i mashin: ucheb posobie*. Moscow, RIOR: INFRA-M.

Bereska, L., Gavves, E. (2024) Mechanistic Interpretability for AI Safety — A Review [online] *Transactions on Machine Learning Research (08/2024)*. Available at: <https://leonardbereska.github.io/blog/2024/mechinterpreview/> (accessed: 03.08.2025).

Bogolyubov, A. N. (1976) *Teoriya mekhanizmov i mashin v istoricheskom razvitii ee idej*. Moscow, Nauka.

Bryant, L. (2012) *Machine oriented ontology* [online] Available at: <https://larvalsubjects.wordpress.com/2012/06/21/machine-oriented-ontology-towards-a-pan-mechanism/> (accessed: 03.08.2025).

Deleuze, G., Guattari, F. (2007) *Anti-Oedipus: Capitalism and Schizophrenia*, Yekaterinburg, U-Factoria. 672 p.

Deleuze, G., Guattari, F. (2009) Balance Sheet for Desiring Machines. In: Guattari, F., Lotringer, S. *Chaosophy: texts and interviews, 1972–1977*. Los Angeles, CA: Semiotext(e). Pp. 90–115.

Engelmeir, P. K. (1887) *Ekonomicheskoe znachenie sovremennoj tekhniki*. Moscow, Russkaya tipolitografiya [online] Available at: [https://archive.org/details/ekonomicheskoe\\_znachenie\\_sovremennoy\\_tekhniki/page/2/mode/2up](https://archive.org/details/ekonomicheskoe_znachenie_sovremennoy_tekhniki/page/2/mode/2up) (accessed: 03.08.2025).

Hardt, M. (1993) *Gilles Deleuze: an Apprenticeship in Philosophy*. Minneapolis, University of Minnesota Press.

Kubryakova, E. S. (2001) O kognitivnoj lingvistike i semantike termina. *Vestnik VGU. Seriya: Lingvistika i mezhkul'turnaya kommunikaciya*, no. 1, pp. 4–10.

Lafitte, J. (1972) *Reflections on the Science of Machines*. London, Mechanology Press.

Land, N. (2020) *Kibergotika*. In: Land, N. *Sochineniya*. In 2 vols. Vol. 1. Perm : HylePress, 2020. Pp. 79–106 (In Russ.).

Le Corbusier (2017) *Novyi dukh v arkhitekture*. Moscow, Strelka Press Publ. (In Russ.).

Malyshev, A. P. (1923) Analiz i sintez mekhanizmov s točki zreniya ih struktury. *Izvestiya of the Tomsk Technological Institute*, vol. 44.

Mandelbrot, B. (2004) *Fraktaly, sluchaj i finansy*. Moscow — Izhevsk, SIC “Regular and chaotic dynamics”.

Menabrea, L. F. (1842) Invented by Charles Babbage. With notes upon the Memoir by the Translator Ada Augusta, Countess of Lovelace from the *Bibliothèque Universelle de Genève*, October, 1842, no. 82. [online] Available at: <https://fourmilab.ch/babbage/sketch.html> (accessed: 03.08.2025).

Moiseev, D. V., Korzhinski, A. (2025) *Nauka i zdravomyslie — 2025. Obnovlenie metodologii obschbej semantiki v XXI veke. Izdatelskiye resheniya* [online] Available at: [https://books.yandex.ru/books/sDFrDlzi/read-online?with\\_payment\\_widget=true](https://books.yandex.ru/books/sDFrDlzi/read-online?with_payment_widget=true) (accessed: 03.08.2025).

Mumford, L. (2001) *Mif mashiny. Tekhnika i razvitie chelovechestva*. Moscow, Logos.

Nikolin, V. V. (2002) *Mashinnost' kak vseobschbij princip vosproizvodstva: Ontologicheskij aspekt ekspansii mashiny v nemashinnye sfery bytiya: Abstract of the dis. ... Dr. of Philosophy*. Omsk : OmGPU.

Obolkina, S. (2025) Kratkij ocherk mashinizma. *Idei i idealy = Ideas and Ideals, vol. 17, iss. 3, pt. 1, pp. 124–143*.

Pearl, J. (2009) *Causality*. Cambridge University Press.

Podoroga, V. (2010) Homo ex machina. Avant-garde and Its Machines. *Aesthetics of a New Form, Logos, no. 1, pp. 22–50*.

Ryklin, M. (2002) *Mashiny` zbelaniya i prosto mashiny* [online] Available at: <https://schizopromenade.blogspot.com/2014/07/guattari.html> (date accessed: 13.12.2024).

Simondon, G. (1958) *Du mode d'existence des objets techniques*. Paris, Aubier.

Stepanov, M. (2010) Apparatoye ponimaniye media. *Bulletin of the Leningrad State University named after A. S. Pushkin*, no. 3, pp. 194–199.

Toward Transparent AI: A Survey on Interpreting the Inner Structures of Deep Neural Networks (2023) / T. Räuber, A. Ho, S. Casper and D. Hadfield-Menell. *Raleigh*, 8–10 February, pp. 464–483.

Wiener, N. (1996) *Tvorec i robot. Obsuzhdenie nekotoryh problem, v kotoryh kibernetika stalkivaetsya s religiej*. Moscow, Progress.

Zaripov, Ya. V. (2015) Machine: analiz semantiki v diahronicheskoy perspective. In: *Proceedings of NEANIAI — All-Russian School of Young Scientists on Byzantine Studies* [online] Available at: <https://www.hse.ru/data/2015/10/03/1077230538/Зарипов%20Я.В.pdf> (accessed: 03.08.2025).

Zednik, C. (2021) Solving the Black Box Problem: A Normative Framework for Explainable Artificial Intelligence. *Philosophy & Technology*, no. 34, pp. 265–288.

*Submission date: 10.01.2026.*

Оболкина Светлана Викторовна — кандидат философских наук, старший научный сотрудник сектора истории и философии науки Института философии и права Уральского отделения РАН. Адрес: 620137, Российская Федерация, г. Екатеринбург, ул. Софьи Ковалевской, д. 16. Тел.: +7 (343) 362-34-07. Эл. адрес: [obol2007@mail.ru](mailto:obol2007@mail.ru)

Obolkina Svetlana Viktorovna, Candidate of Philosophy, Senior Researcher, Sector of the History and Philosophy of Science, Institute of Philosophy and Law, Ural Branch, Russian Academy of Sciences. Postal address: 16, Sofyi Kovalevskoi St., Yekaterinburg, Russian Federation, 620137. Tel.: +7 (343) 362-34-07. E-mail: [obol2007@mail.ru](mailto:obol2007@mail.ru)