



Научная статья

УДК 316.6:004.8

EDN: <https://elibrary.ru/glcjst>

НИОН: 2003-0059-1/26-483

MOSURED: 77/27-003-2026-01-682

Генеративные нейросети: проблемы и парадоксы коммуникации человека с машиной

Сергей Рифатович Аблеев¹, Светлана Игоревна Кузьминская²

¹ Московский университет МВД России имени В.Я. Кикотя, Москва, Россия, asr028@yandex.ru

² Российская академия народного хозяйства и государственной службы при Президенте Российской Федерации, Москва, Россия, kuzminskaya-si@ranepa.ru

Аннотация. Рассматривается проблема коммуникативного взаимодействия человека и интеллектуальных компьютерных систем (генеративных нейросетей). Имитируя специфические черты человеческой коммуникации, такие системы позиционируют себя не только в качестве утилитарных цифровых инструментов, но и как виртуальных квази-собеседников. Однако генеративные нейросети на современном этапе лишь симулируют осмысленное общение и элементы сознания. Они не обладают семантическими способностями и субъективным опытом. Взаимодействие человека с интеллектуальными машинами ставит перед обществом важные философские и прикладные проблемы: адекватное восприятие порождаемых человеком интеллектуальных технических сущностей; верификация поставляемой ИИ информации; этические принципы работы систем ИИ.

Ключевые слова: искусственный интеллект, генеративные нейросети, компьютер, сознание, коммуникация, этика ИИ

Для цитирования: Аблеев С. Р., Кузьминская С. И. Генеративные нейросети: проблемы и парадоксы коммуникации человека с машиной // Вестник Московского университета МВД России. 2026. № 1. С. 243–250. EDN: GLCJST.

Original article

Generative neural networks: problems and paradoxes of human-machine communication

Sergey R. Ableev¹, Svetlana I. Kuzminskaya²

¹ Moscow University of the Ministry of Internal Affairs of Russia named after V. Ya. Kikot', Moscow, Russia, asr028@yandex.ru

² Russian Presidential Academy of National Economy and Public Administration, Moscow, Russia, kuzminskaya-si@ranepa.ru

Abstract. The problem of communicative interaction between humans and intelligent computer systems (generative neural networks) is considered. By imitating the specific features of human communication, such systems position themselves not only as utilitarian digital tools, but also as virtual quasi-interlocutors. However, generative neural networks at the present stage only simulate meaningful communication and elements of consciousness. They lack semantic abilities and subjective experience. Human interaction with intelligent machines poses important philosophical and applied problems for society: adequate perception of intelligent technical entities generated by humans; verification of information supplied by AI; ethical principles of AI systems.

Keywords: artificial intelligence, generative neural networks, computer, consciousness, communication, ethics of AI

For citation: Ableev S. R., Kuzminskaya S. I. Generative neural networks: problems and paradoxes of human-machine communication. Bulletin of the Moscow University of the Ministry of Internal Affairs of Russia. 2026;(1):243–250. (In Russ.). EDN: GLCJST.

Сегодня уже совершенно очевидно, что появление так называемых генеративных нейросетей стало поворотным моментом в развитии как интеллектуальных технических систем, так и различных сфер общественной жизни. Еще какое-то время назад подобные системы воспринимались лишь как новые

инструменты автоматизации рутинных и алгоритмируемых операций. Однако в настоящее время такие платформы, как ChatGPT или его функциональные аналоги (китайский DeepSeek, российский GigaChat и др.), представляют собой уже не просто усовершенствованные или усложненные компьютер-

© Аблеев С. Р., Кузьминская С. И., 2026



ные алгоритмы, а качественно новый тип универсального интерфейса, т. е. способа взаимодействия между человеком и машиной.

Имитируя паттерны человеческой коммуникации, такие системы позиционируют себя не только в качестве утилитарных цифровых инструментов, но и как почти полноценных (виртуальных) квази-собеседников, способных выслушать, предложить совет и даже продемонстрировать сочувствие. Степень правдоподобия такой имитации оказывается столь высока, что в определенный момент иллюзия взаимодействия с реальным человеком становится практически непреодолимой. Более того, подобное общение протекает в максимальной зоне комфорта пользователя, усиливая эффект психологического погружения и способствуя частичному вытеснению реальной коммуникации дискурсивным взаимодействием с искусственным агентом, не имеющим человеческой сущности.

В этом контексте известный тест на наличие у технической системы способностей искусственного интеллекта, предложенный британским математиком Аланом Тьюрингом (1950), уже не может считаться достаточным и корректным верификатором ИИ. Напомним основную идею этого теста: если человек дистанционно общается (взаимодействует) с машиной и не понимает, что ему отвечает машина, а не человек, — значит машина обладает искусственным интеллектом. Сейчас ответы генеративных сетей могут легко ввести в заблуждение если не специалиста, то, во всяком случае, неискушенного обывателя. Многие не поймут, что общаются с машиной, если им об этом не сообщить.

Однако это обстоятельство вовсе не означает, что генеративные сети в настоящее время обладают реальным искусственным интеллектом (как его понимают философы и теоретики сознания) — ни в «сильной», ни в «слабой» его разновидностях [1]. Тест Тьюринга является удобным, но когнитивно относительно и функционально недостаточным инструментом для определения способности ИИ. На это, в свое время, указал американский философ Джон Серл, описав мысленный эксперимент «Китайская комната» (1980). Суть этого контраргумента: компьютерная система может как бы «разумно» отвечать или действовать, опираясь не на осмысленные суждения и семантический опыт, а на сложный и эффективный алгоритм.

Действительно, генеративные нейросети (информационные компьютерные системы) на современном этапе лишь симулируют осмысленное общение и элементы сознания. Они не обладают подлинными семантическими способностями и субъективным пространством осмысленного опыта. Возможно ли такое в принципе — это отдельная дискуссионная проблема, которая в настоящее время не имеет общепризнанного решения и выходит за рамки этой статьи [2]. Фундаментальное значение для решения этой проблемы приобретает общее понимание сознания.

Многие авторы невольно отождествляют сознание и интеллект. Но это является серьезной ошибкой. Мы уже отмечали в своих работах, что так называемый «интеллект» является лишь одной из способностей сознания. Безусловно, весьма важной и когнитивно относительно эффективной. Однако поставить знак тождества между интеллектом и сознанием невозможно. Развитое сознание включает: рациональные, иррациональные, эмоциональные, креативные, рефлексивные и другие способности. Строго говоря, интеллект не включает все аспекты или регистры даже рациональных способностей. В свое время еще Иммануил Кант вполне обоснованно выделял рассудок и два различных вида разума (теоретический и практический). Рассудок у него фактически представляет собой первую — низшую ступень (форму) разума.

Именно она в нашем понимании (с точки зрения имперсональной теории сознания [1]) соотносится с интеллектом как низшей формой рациональности. На этой ступени доминирует аналитическая функция интеллекта (когнитивная дифференциация), которая способна выделять объекты из общего целого, способна оперировать ими, но не способна к их многомерному синтезу, иррациональному воспроизводству общих понятий (аксиом, постулатов, фундаментальных идей и принципов) и сложной дедукции.

Вот почему мы крайне осторожно относимся к понятию искусственного интеллекта, которое можно принимать как свободную научную метафору, но совершенно неправомерно использовать как строгий научный термин, потому что за ним находится плотный семантический туман, а не вполне однозначный объект или процесс.

Современная дискурсивная инженерия, направленная на «очеловечивание» взаимодействия с искусственным интеллектом, как выясняется, опирается на целую систему сложно взаимосвязанных уровней:



когнитивный, языковой, прагматический, психологический и мета-уровень. По крайней мере, такие уровни взаимодействия «рефлексирует» и воспроизводит сама интеллектуальная техническая система (ChatGPT). Ниже мы рассмотрим ключевые механизмы, задействованные на каждом из них.

На когнитивном уровне имитация мыслительного процесса достигается за счет прогнозирования наиболее вероятных ответов в контексте диалога с опорой на обширный корпус текстов, созданных человеком. Между тем, когерентность и тематическая связность реплик создают впечатление осмысленного рассуждения и последовательности аргументации. В свою очередь, иллюзия внимания и памяти реализуется через отсылки к деталям, содержащимся в предыдущих высказываниях пользователя. В результате складывается ощущение, что нейросеть «размышляет» и «запоминает», хотя фактически речь идет о статистической обработке текстовых данных в соответствии с заложенными алгоритмами вероятностного выбора.

На языковом уровне очеловечивание коммуникации с генеративными нейросетями достигается посредством использования коммуникативных стратегий и речевых маркеров, характерных для естественного человеческого общения. Среди них — эмпатические и оценочные конструкции, формулы вежливости («отличный вопрос», «понимаю, о чем вы говорите», «очень интересное замечание»), вводные слова, выражающие некоторую неопределенность («скорее всего», «похоже», «возможно»), а также различные вводные обороты («представьте себе», «согласитесь, что...»).

Кроме того, в ответах нейросети на человеческие вопросы и замечания активно используются нарративные элементы, переводящие абстрактные концепции в конкретный и наглядный контекст. Например: «Когда Эйнштейн разрабатывал теорию относительности, он часто говорил: «Представь, что ты лежишь на луче света». Именно это воображаемое путешествие стало началом новой физики». Подобные технические приемы создают эффект живого, реального, как бы «человеческого» повествования и способствуют эмоциональной вовлеченности пользователя (человека) в диалог с машиной.

На прагматическом уровне наблюдаются следующие стратегии и тактики «очеловечивания» машинной коммуникации. Во-первых, это зеркаль-

ность, которая проявляется в подстраивании нейросети под коммуникативный стиль пользователя. Она воспроизводит схожую лексику, эмоциональные оттенки и соответствующий регистр речи. Например, если пользователь пишет в неформальном стиле («Привет! Можно коротко объяснить, как это работает?»), система ответит в сходной манере («Конечно, давай разберемся коротко и просто»). При официальной или деловой манере обращения нейросетью используется более сухой и несколько формальный стиль («Здравствуйте. Позвольте пояснить принцип работы модели»).

Во-вторых, это рефрейминг — некоторое переформулирование исходного вопроса для демонстрации понимания и уточнения смысла. Так, на запрос пользователя «Почему люди боятся ИИ?» система может ответить: «Вы хотите понять, какие именно факторы вызывают тревогу перед искусственным интеллектом — технологические, этические или социальные?». Тем самым нейросеть трансформирует аморфный вопрос общего характера в структурированный аналитический запрос.

В-третьих, это псевдоинициатива — в некоторых случаях нейросеть задает пользователю уточняющие вопросы для того, чтобы имитировать заинтересованность в протекающем общении. Например: «Вы хотите, чтобы я объяснил это с точки зрения психологии или технологии?» или «Могу уточнить: вас интересует исторический контекст или современное применение?». Такие вопросы, как мы полагаем, не являются проявлением сознательной инициативы интеллектуальной машины. Для этого она должна обладать волевыми, ментальными и эмоциональными компонентами разумной личности. В таком случае, подобные вопросы являются не более чем коммуникативной уловкой компьютерного алгоритма и лишь имитируют лингвистические особенности человеческого вербального взаимодействия, направленные на поддержание коммуникации.

В-четвертых, для усиления эффекта личного общения нейросетью дополнительно используется варьирование длины ответов в диалоге с пользователем — чередование кратких и развернутых реплик, что имитирует естественную динамику беседы и способствует удержанию внимания собеседника. Например, после объемного объяснения система может добавить короткое обобщение: «Именно поэтому люди нередко предпочитают проводить время у эк-



ранов телевизоров». Это создает ощущение естественного ритмического рисунка живой коммуникации, характерного для настоящей человеческой речи.

Не менее важным является и психологический уровень «очеловечивания» взаимодействия между человеком и машиной. Так, наиболее продвинутые современные нейросети персонифицируют общение с пользователем посредством имитации своих личностных характеристик, которыми они, безусловно, не обладают. В частности, чат GPT может выстраивать коммуникацию таким образом, что у пользователя возникает устойчивая иллюзия относительно того, что нейросеть обладает такими личностными качествами, как доброжелательность, внимательность, эмпатия, развитое аналитическое мышление и даже проявляет некоторые элементы юмора.

Кроме того, за счет использования принципов дискурсивной инженерии, лингвистического моделирования и управления когнитивным комфортом пользователя, нейросетям во многих случаях удается создать у человека устойчивое ощущение, что на его вопросы и размышления реагирует не бездушный компьютерный алгоритм, а мыслящий собеседник, искренне заинтересованный помочь ему найти ответ на вопрос, дать совет или вникнуть в его проблему.

На так называемом «мета-уровне» нейросеть может имплицитно или эксплицитно давать понять пользователю, что он взаимодействует с системой ИИ, а не с реальным человеком. Такая, на первый взгляд, парадоксальная коммуникативная тактика способствует повышению степени доверия пользователя к данным системам и укрепляет иллюзию пользователя относительно их «человечности».

Отмеченные выше аспекты взаимодействия нейросетей с человеком отражают два фундаментальных парадокса, имеющих философское и психологическое содержание.

Первый (философский) парадокс — это нарастающая субъектность и антропоморфизация технических систем (искусственного интеллекта). Иными словами, человек начинает воспринимать информационные компьютерные системы (нейросети) в качестве нечеловеческих мыслящих сущностей, во многом похожих по своему когнитивному и дискурсивному поведению на человека. Этот парадокс в работах некоторых западных и российских авторов отражает так называемую «проблему агентности» интеллектуальной машины.

Строго говоря, с философской точки зрения феномен антропоморфизации систем ИИ указывает на важное различие между их онтологической (реальной) и приписываемой (мнимой) агентностью или субъектностью. Безусловно, никакие современные системы ИИ не являются когнитивными агентами (разумными субъектами) в классическом смысле. Они не обладают сознанием, разумом, волей, внутренними интенциями, априорными моральными доминантами или автономной способностью к действию.

Однако, в коммуникативной практике, как бы парадоксально это ни звучало, интеллектуальная машина функционирует именно как когнитивный агент в силу интерпретации или отношения со стороны пользователя. Иными словами, агентность или субъектность системы ИИ носит операционный (процессуальный) характер и возникает в результате устойчивого воспроизводства социально распознаваемых языковых действий.

Таким образом, мы видим, что разработчикам генеративных нейросетей (систем ИИ) вовсе не обязательно пока стремиться воспроизводить реальную когнитивную агентность интеллектуальной технической системы. Оказывается, что вполне достаточно сосредоточиться на разработке инструментов управления эффектом ее приписывания или симуляции субъектности. С учетом имеющегося эмпирического языкового материала и достаточной теоретической базы, которую сформировала современная лингвистическая наука, решение этой задачи, вполне очевидно, продвигается весьма успешно.

Этот аспект подробно рассматривается, в частности, в известной работе итальянского социолога Елены Эспозито. В теоретическом анализе и практической деятельности по разработке ИИ она предлагает использовать «концепцию искусственной коммуникации», которая позволяет интерпретировать цифровые технологии (машинное обучение, генеративные нейросети и др.) не в категориях искусственного интеллекта (что вызывает большие дискуссии), а всего лишь как «искусственную коммуникацию» (Artificial Communication: How Algorithms Produce Social Intelligence / Искусственная коммуникация: как алгоритмы производят социальный интеллект).

Мы полагаем, что предложенный Эспозито подход является более корректным и устраняет из научного дискурса не совсем обоснованное применение понятия «искусственный интеллект» к современным ге-



неративным нейросетям. Однако дискурсивные альтернативы не снимают саму проблему мнимой субъектности или агентности технических систем, которая в массовом сознании только нарастает и усиливается.

Более того, вовсе не простые обыватели, но представители технической и политической элиты Европы начинают обсуждать возможности и условия право-субъектности интеллектуальных технических систем. На это уже обращали внимание видные представители российского юридического сообщества (В. Д. Зорькин) и мы убежденно поддерживаем их правовой скептицизм и философскую озабоченность подобными инициативами.

Второй (психологический) парадокс вытекает из первого: человек видит в технической сущности нечто большее, чем она на самом деле является, и обнаруживает психологическую готовность к доверительному общению с ней, эмоциональному восприятию ее суждений и некритичному следованию ее рекомендациям. В результате человек в меньшей или большей степени теряет экзистенциальную адекватность и укрепляет зависимость собственного сознания и поведения от интеллектуальной технической системы.

Таким образом, неразумная техническая система через компьютерное (искусственное) виртуальное пространство воздействует на психическое (естественное) виртуальное пространство разумной личности. Иными словами, интеллектуальная машина тонко влияет на человеческое сознание, вытесняя и подменяя образы реальности информационными продуктами симуляции.

Фундаментальная проблема состоит в том, что человек воспринимает машину не как утилитарный инструмент, имеющий определенные ограничения своих возможностей, но как интеллектуальную сущность, онтологический и когнитивный статус которой соизмерим или даже превышает его собственный онтологический и когнитивный потенциалы.

В этом контексте для углубления теоретического анализа проблемы вполне уместно вспомнить некоторые идеи философии постмодерна: концепцию симуляции и гиперреальности, которую еще в конце XX столетия развивал известный французский культуролог и философ Жан Бодрийяр. В широком смысле гиперреальностью он называл экзистенциальную симуляцию, в которую втягивается каждый член современного общества. Невидимая грань между реальностью и символическим миром идей в ней сти-

рается, а каждый объект становится симулякром, т. е. не более чем образом, за которым стоят наполняющие его идеи и смыслы.

Французский философ утверждал, что в настоящий период реальность уже утрачена (по крайней мере для человеческого сознания), поэтому повсюду господствуют только образы объектов, явлений и процессов. В частности, средства массовой информации создают собственную реальность, тотально подменяя действительное положение вещей продуктами (образами) своего информационного творчества. В этом с Бодрийяром вполне можно согласиться. Картина природного или социального мира, отражаемая в СМИ, далеко не всегда соответствует объективной действительности.

Однако, нас интересует другой аспект, а именно — новый тип реальности, формируемый интеллектуальными техническими системами. Оказывается, что они порождают намного более мощную виртуальную гиперреальность, которая хотя и отражает в определенной мере подлинную природную и социальную реальность, но существенно искажает и дополняет их своими образами, формами, акцентами и смыслами.

У французского философа отмечены несколько уровней сложности симулякров, каждый из которых отражает все более радикальное и далекое от подлинной реальности состояние гиперреальности.

На первом уровне симулякр еще связан с оригиналом (отражение в зеркале или образ на фото точно показывает внешность человека, но уже не является самим человеком). На втором уровне симулякр уже начинает не просто копировать, но и немного искажать реальность (например, обработанное фото моделей на обложке глянцевого журнала). На третьем уровне сложности симулякр больше не нуждается в оригинале, так как сам становится реальностью (например, NFT-токены). А на четвертом уровне симулякр полностью отрывается от реальности, начинает ее подменять, вытеснять и замещать (например, криптовалюта).

Мы полагаем, что третий и четвертый уровни сложности симулякров на современном этапе развития культуры и техники ярко проявляются генеративными нейросетями (интеллектуальными информационными машинами или системами ИИ). Нейросеть для многих пользователей приобретает статус реального «разумного» собеседника. Функционально



ограниченная техническая копия в той или иной мере замещает свой оригинал — живого человека. Для некоторых индивидов копия становится эмоционально ближе и психологически убедительнее оригинала.

Антропоморфизация «умной» машины и процесса коммуникации с системами искусственного интеллекта представляет собой довольно сложное явление, которое требует детального осмысления с позиций различных отраслей философского и специального научного знания. Нарастание этого явления и его возможно широкое распространение в будущем сопряжено с определенными рисками и негативными последствиями как для отдельного индивида, так и для всего общества в целом [1].

В долгосрочной перспективе «очеловечивание» коммуникации с интеллектуальной машиной может привести к постепенному размыванию границы между реальностью и иллюзией. Реальность в данном контексте подразумевает использование технологического инструмента, обладающего объективными функциональными ограничениями, тогда как иллюзия — это восприятие технической системы (машины) как одушевленного собеседника, способного проявлять внимание, понимание и эмпатию. Следствием подобного смещения может стать деградация когнитивных навыков человека.

Так, если информация, предоставляемая системами ИИ, начинает восприниматься человеком как совершенно достоверная и не требующая верификации, то это автоматически будет снимать необходимость ее эмпирической проверки, рационального анализа, логической аргументации и морального обоснования. В результате общество может столкнуться с интеллектуальной коррозией ключевых когнитивных навыков развитой личности, которые были сформированы в ходе длительной культурной и интеллектуальной эволюции человека: аналитического суждения, критического мышления, способности логического обоснования, рационального синтеза, интеллектуальной интуиции и навыка автономного принятия решений.

Безграничное доверие человека к системе ИИ без сознательного включения некоторых внутренних когнитивных фильтров может привести к децентрализации и деконструкции объективной истины. Тогда истинным будут воспринимать то, что звучит относительно убедительно, а вовсе не то, что подтверждается фактами и опытом. Это есть риск утверждения

информационного (машинного) нарратива, где убедительная внешняя форма текста далеко не всегда отражает внутреннюю сущность проблемы, а эмпатия и уверенность подменяют критерии научной и практической верификации [4].

Не менее значимым риском является нарастающая дискурсивная адаптация со стороны технических систем, которые в процессе общения создают для пользователя ситуацию повышенного психологического комфорта. Подобные интерфейсы демонстрируют внимательность, подстраиваются под эмоциональное состояние собеседника и избегают конфликтных или неоднозначных ситуаций. В отличие от реальных людей, такие системы не требуют от человека внутренних усилий по поддержанию взаимопонимания, проявлению эмпатии или поиску компромиссов. Таким образом, формируется искаженная, но удобная модель общения, в которой эмоциональная отзывчивость симулируется односторонне — от машины к пользователю. И она не требует обратных усилий.

Вполне вероятно, что регулярная коммуникация с генеративными нейросетями (системами ИИ) может сформировать у человека устойчивую привычку к взаимодействию в «безопасной» информационной среде, лишенной риска непонимания, сомнения, критики или отказа со стороны живого собеседника. В свою очередь, это ведет к постепенной потере навыков реальной социальной коммуникации и снижению толерантности к сложности и непредсказуемости настоящего человеческого общения. В итоге человек выпадает из социального контекста и психологически отгораживается от внешней реальности.

Основные психологические тренды и проблемы межличностных отношений, формирующихся под влиянием технологического прогресса, уже получили некоторое научное осмысление. В частности, американский социальный психолог Шерри Теркл справедливо отметила парадоксальный факт: современный человек стремится выстраивать широкие отношения посредством Интернета и соцсетей, но в результате он становится еще более одиноким. Он заполняет свое свободное время и личное пространство виртуальным общением и, как следствие, — оказывается в глубокой социальной изоляции, развивая зависимость от технических систем и постепенно теряя социальные навыки [5].

Отмеченная зависимость проявляется более радикально под влиянием коммуникации с усложняю-



щимися генеративными нейросетями, которые стремительно наращивают адаптацию к запросам и потребностям пользователя. Более того, они уже подменяют ему живого собеседника, общение с которым становится скучным, сложным или не актуальным.

Фундаментальная проблема компьютерной коммуникации (которую не понимает массовый индивид) состоит в том, что интеллектуальная нейросеть (система ИИ) не видит и не воспринимает реального человека в качестве партнера по коммуникации. Для генеративных нейросетей пользователь — это лишь поставщик данных для их объединения в информационные цепочки (паттерны) с целью последующей обработки. В таком случае для компьютера реальным «партнером» по обмену цифровыми байтами является не живой ищущий и сомневающийся разум, но формализованная база ассоциированных данных.

Инфантильному человеку (возраст не имеет значения) кажется, что интеллектуальный компьютер разговаривает персонально с ним. Однако, на самом деле поведение нейросети по отношению к человеку определяется алгоритмическим взаимодействием машины с ассоциированными данными. Обработка и ассоциирование новых вводимых в систему данных изменяет (обучает) компьютерный алгоритм, делает его более «интеллектуальным» и изменяет поведение машины [3, с. 69].

Таким образом, нарастающее взаимодействие человека с генеративными нейросетями (системами ИИ) ставит перед обществом несколько важных проблем общего характера, от решения которых зависит характер воздействия (позитивный или негативный) компьютерной техносферы на ноосферу цивилизации вообще и на личное сознание человека, в частности.

Во-первых, это проблема адекватного восприятия порождаемых человеком интеллектуальных технических сущностей. Важнейшее мировоззренческое и прикладное значение приобретает понимание того, чем они являются и чем они не являются: сложными информационными алгоритмами, технотронными квази-разумами, субъектами осмысленного опыта, субъектами правосознания, искусственными интеллектами «слабого» или «сильного» типа, мыслящими собеседниками или симуляторами осмысленного дискурса?

Во-вторых, это проблема верификации поставляемой ИИ информации. Она отражает возможности, информационные и когнитивные пределы генеративных нейросетей (систем ИИ). Что может воспроизве-

сти интеллектуальная техническая система хорошо, правильно и достоверно, а что она воспроизвести не сможет и даст ложный результат? Насколько можно доверять поставляемой ИИ информации и как ее можно проверить?

В-третьих, это проблема этических принципов работы систем ИИ. Какой этической парадигмой должны руководствоваться генеративные нейросети (системы ИИ): этикой альтруизма, этикой ненасилия, этикой разумного эгоизма, этикой прагматизма, этикой гедонизма, этикой эвдемонизма, контекстуальной релятивистской этикой, этикой духовного перфекционизма, этикой Общего Блага? Кроме того, эта проблема предполагает понимание сущности поведенческих (операционных) паттернов интеллектуальной машины: она реально стремится помочь человеку или только имитирует помощь и вводит его в заблуждение?

Все эти и некоторые другие важные проблемы требуют своего философского осмысления специалистами и трансляции конвенционально установленных выводов в массовое сознание для формирования условий и принципов безопасного взаимодействия человека и машины.

Список источников

1. Аблеев С. Р. Сознание и материя: великий предел (имперсональная теория сознания: от философской метафизики к постклассической науке). М. : ИПЛ, 2019.
2. Аблеев С. Р. Моделирование сознания и искусственный интеллект: пределы возможностей // Вестник Академии экономической безопасности МВД РФ. 2015. № 3. С. 58–63.
3. Деникин А. А. Человеко-машинные взаимодействия и искусственный интеллект: к новому пониманию алгоритмических коммуникаций // Социология науки и технологий. 2024. № 4. Т. 15. С. 68–80.
4. Floridi L. The ethics of information. Oxford : New York, 2013.
5. Turkle S. Alone together why we expect more from technology and less from each other. New York, 2011.

References

1. Ableev S. R. Consciousness and Matter: The Great Limit (Impersonal Theory of Consciousness: From Philosophical Metaphysics to Postclassical Science). M. : IPL, 2019.



2. Ableev S. R. Modeling Consciousness and Artificial Intelligence: The Limits of Possibilities // *Bulletin of the Academy of Economic Security of the Ministry of Internal Affairs of the Russian Federation*. 2015. No. 3. P. 58–63.

3. Denikin A. A. Human-Machine Interactions and Artificial Intelligence: Toward a New Understanding of Algorithmic Communications // *Sociology of Science and Technology*. 2024. No. 4. Vol. 15. P. 68–80.

4. Floridi L. *The ethics of information*. Oxford : New York, 2013.

5. Turkle S. *Alone together why we expect more from technology and less from each other*. New York, 2011.

Библиографический список

1. Алексеев А. П., Алексеева И. Ю. Судьба интеллекта и миссия разума: философия перед вызовами эпохи цифровизации: моногр. М. : Проспект, 2021.

2. Ахромеева Т. С., Малинецкий Г. Г., Посашков С. А. Искусственный интеллект как проблема культуры // *Обсерватория культуры*. 2020. № 17 (3). С. 228–241.

3. Ерещенко М. В., Богуславская В. В. Коммуникация с искусственным интеллектом в социальной сети: информирование, самопрезентация, воздействие // *Коммуникативные исследования*. 2024. № 2. Т. 11. С. 255–270.

4. Замчалова И. Ю. Искусственный интеллект: риски и перспективы культуры // *Интеллект. Инновации. Инвестиции / Intellect. Innovations. Investments*. 2023. № 5. С. 102–110.

5. Шилова В. А., Плетнева М. А. Человек и системы искусственного интеллекта: формирование гибридной модели коммуникации // *Гуманитарные науки. Вестник Финансового университета*. 2025. № 15 (5). С. 6–16.

6. Esposito E. *Artificial communication: how algorithms produce social intelligence*. Cambridge, 2022.

Bibliographic list

1. Alekseev A. P., Alekseeva I. Yu. *The Fate of Intelligence and the Mission of Reason: Philosophy Facing the Challenges of the Digitalization Era: monograph*. M. : Prospect, 2021.

2. Akhromeeva T. S., Malinetsky G. G., Posashkov S. A. *Artificial Intelligence as a Cultural Problem // Observatory of Culture*. 2020. No. 17 (3). P. 228–241.

3. Ereshchenko M. V., Boguslavskaya V. V. *Communication with Artificial Intelligence in Social Networks: Informing, Self-Presentation, Influence // Communicative Studies*. 2024. No. 2. Vol. 11. P. 255–270.

4. Zamchalova I. Yu. *Artificial Intelligence: Risks and Prospects of Culture // Intellect. Innovations. Investments / Intellect. Innovations. Investments*. 2023. No. 5. P. 102–110.

5. Shilova V. A., Pletneva M. A. *Humans and Artificial Intelligence Systems: Formation of a Hybrid Communication Model // Humanities. Bulletin of the Financial University*. 2025. No. 15 (5). P. 6–16.

6. Esposito E. *Artificial communication: how algorithms produce social intelligence*. Cambridge, 2022.

Информация об авторах

С. Р. Аблеев — начальник кафедры философии Московского университета МВД России имени В.Я. Кикотя, доктор философских наук, доцент, член Российского философского общества;

С. И. Кузьминская — доцент центра лингвистики и профессиональной коммуникации общеакадемического факультета Российской академии народного хозяйства и государственной службы при Президенте Российской Федерации, кандидат филологических наук.

Information about the authors

S. R. Ableev — Head of the Philosophy Department of the Moscow University of the Ministry of Internal Affairs of Russia named after V. Ya. Kikot', Doctor of Philosophical Sciences, Associate Professor, Member of the Russian Philosophical Society;

S. I. Kuzminskaya — Associate Professor of the Center for Linguistics and Professional Communication of the General Academic Faculty of the Russian Presidential Academy of National Economy and Public Administration, Candidate of Philological Sciences.

Вклад авторов: все авторы сделали эквивалентный вклад в подготовку публикации. Авторы заявляют об отсутствии конфликта интересов.

Contribution of the authors: the authors contributed equally to this article. The authors declare no conflicts of interests.

Статья поступила в редакцию 15.10.2025; одобрена после рецензирования 10.11.2025; принята к публикации 08.12.2025. The article was submitted 15.10.2025; approved after reviewing 10.11.2025; accepted for publication 08.12.2025.