

ПАРТНЕРСКИЕ ОТНОШЕНИЯ ЧЕЛОВЕКА И ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: МИФ ИЛИ РЕАЛЬНОСТЬ В СОВРЕМЕННОМ МИРЕ

Аверкин А.Н., канд. физ.-мат. наук, доцент, ведущий научный сотрудник Федерального исследовательского центра «Информатика и управление» Российской академии наук

Воронина Н.Ф., канд. социолог. наук, доцент, профессор Академии военных наук, доцент кафедры «Экономика и управление авиационной промышленности» Московского авиационного института (МАИ)

Пелипас О.В., руководитель межведомственной рабочей группы «Технические средства защиты» комитета Торгово-промышленной палаты Российской Федерации по безопасности предпринимательской деятельности, аспирант, начальник отдела информационной безопасности Российского института стандартизации

В статье авторы исследуют возможные формы взаимодействия человека и искусственного интеллекта (ИИ). Рассматриваются вопросы дополнения человека ИИ, процессы взаимного влияния друг на друга и возникающие в связи с этим эффекты и угрозы, выявляются пути взаимодействия, в том числе с учетом принципов построения генеративных нейросетей. Особое внимание уделено выявлению особенности восприятия нейросетью себя и признания своей «разумности», что позволяет авторам сделать допущение о взаимодействии человека с ИИ на основе партнерских отношений. В статье приведены результаты интервьюирования ведущих генеративных нейросетей посредством чат-ботов, и на основе полученного материала авторами сделаны соответствующие выводы о достижимости формирования партнерских отношений в системе «человек – ИИ».

Ключевые слова: искусственный интеллект, нейросеть, человек, разум, самоидентификация.

Цитирование: Аверкин А.Н., Воронина Н.Ф., Пелипас О.В. Партнерские отношения человека и искусственного интеллекта: миф или реальность в современном мире // Информационно-экономические аспекты стандартизации и технического регулирования. 2025. № 2(83). С. 44–51.

ВВЕДЕНИЕ

Партнерские отношения человека и искусственного интеллекта в современном мире могут стать основой для фундаментальных изменений в способах работы и взаимодействия в системе «человек – ИИ».

Интеграция ИИ в социально-экономические, технологические и иные процессы в активном взаимодействии с человеком приводит к необходимости осознания места ИИ в структуре взаимоотношений «человек – ИИ». В 2024 году ChatGPT успешно прошла тест Тьюринга. В 2025 году ей на смену пришла версия o3, создатели которой заявляют о создании AGI в ближайшие годы. Использование чат-ботов в разных формах, в том числе в общении с клиентами, при ведении переговоров, внедрились в свою деятельность 80% компаний в мире [1].

При формировании партнерских отношений необходимо учитывать технические, этические, социальные, культурные и экономические аспекты. Процесс разработки и внедре-

ния ИИ направлен на расширение возможностей человека, а не на подмену, чтобы преимущества от использования ИИ были доступны всем.

Европарламент разработал всеобъемлющую европейскую промышленную политику в области искусственного интеллекта и робототехники (A comprehensive European industrial policy on artificial intelligence and robotics), в рамках которой с 2016 года разрабатывался проект о признании роботов электронными личностями и наделении их определенными правами и обязанностями в обществе, что позволяет рассматривать ИИ с точки зрения вовлечения в партнерские отношения, а не использования ИИ для замены человека в различных отраслях экономики.

В РФ активно обсуждаются правосубъектность роботизированных систем: «... в рамках все более широкого внедрения роботизированных систем право, выступая в роли инструмента стратегического управления в обществе, имеет свои отличительные особенности, выделяющие его среди других средств управления» [2].

Множественные исследования показывают, что при решении одних и тех же задач ИИ и человек затрачивают разное время, но отсутствуют исчерпывающие данные, направленные на характеристику факторов риска, возникающих для человека в результате доверия результатам ИИ, что определяет целесообразность темы исследования.

МЕТОДЫ И МЕТОДОЛОГИЯ

Целью исследования является выявление возможности партнерских отношений в системе «человек – ИИ». Ключевым критерием является допущение, что ИИ может стать самостоятельным субъектом для реализации функций партнера. Признание нейросетью своей разумности могло бы характеризовать готовность человека, на обучающей выборке которого обучалась нейросеть, к партнерским отношениям с ИИ.

Задачами исследования является интервьюирование периферических на момент исследования генеративных нейросетей различных производителей посредством чат-ботов с целью получения признания от них в равенстве с человеком (самостоятельный субъект), т.е. в обладании разумности.

Признавая, что генеративная нейросеть не является общим искусственным интеллектом (Artificial general intelligence, AGI), полноценным ИИ, авторы все же используют этот вид нейросетей в качестве прообраза AGI и в дальнейшем будут называть интервьюируемые нейросети «ИИ».

В исследовании учитывалось, что исследуемые генеративные нейросети являются нейросетями без проверки подлинности, в связи с этим, авторы трактовали полученные результаты исходя из известных принципов построения нейросетей.

В процессе проведения интервью с большими языковыми моделями (LLM) авторы стремились к признанию ИИ самого себя разумным в ходе одного сеанса разговора, и оценивали потребовавшееся для этого время.

Также авторы стремились в интервьюировании оценить «креативность» ИИ в тематике диалога, а именно его способность генерировать самостоятельные или неожиданные, перспективные, с их точки зрения темы разговора *или* идеи.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ

В настоящее время не существует единой парадигмы в отношении ИИ.

ГОСТ Р 59277–2020 определяет ИИ (artificial intelligence) как «комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая самообучение, поиск решений без заранее заданного алгоритма и достижение инсайта) и получать при выполнении конкретных практически значимых задач обработки дан-

ных результаты, сопоставимые, как минимум, с результатами интеллектуальной деятельности человека».

В трактовке Европарламента ИИ представляет собой следующее «... система, основанная на машинах, которая предназначена для работы с различным уровнем автономии и которая может для достижения явных или неявных целей, генерировать результаты, такие как прогнозы, рекомендации или решения, которые влияют на физическую или виртуальную среду» («... a machine-based system that is designed to operate with varying levels of autonomy and that can, for explicit or implicit objectives, generate outputs such as predictions, recommendations, or decisions, that influence physical or virtual environments»).

В обоих определениях говорится или об имитации когнитивных функций человека, но «не хуже чем...», или об автономии, генерировании результатов, влияющих на физическую и виртуальную среду. Учитывая, что деятельность человека также влияет на физическую и виртуальную среду, то уместно задаться вопросом: почему мы заведомо отказываем машинам в равенстве, если последствия существования человека и машины схожи?

ИСТОРИЧЕСКАЯ РЕТРОСПЕКТИВА

Несмотря на многовековую историю изучения разума человека, способов мышления и обработки информации, с античных времен до современного периода, отсутствует единое определение мышления, разума и составляющих их процессов, критериев (и их необходимости) разделения разума на искусственный и естественного происхождения.

Исторически понятие разума претерпевало изменения на протяжении тысячелетий. Гераклит одним из первых стал разделять способность мышления и накопления знаний от разума. Признавая способность мышления и возможность познания общими для всех людей, он полагал, что далеко не всех людей эта возможность приводит к действительному познанию: сколь не было бы обширно усвоенное знание, оно не может заменить подвижность ума, говорил он. Кроме того, он же возводил понятие ratio (с лат. считать, думать) к греческому logos (речь, слово, мера), утверждая его гармонизирующую, умиротворяющую роль, таким образом сопоставляя с разумом свойство гармонизации, умиротворения его окружающего естества [3, с. 73].

В последующих суждениях исследователей обсуждающиеся понятия претерпели значительное усложнение в интерпретации разума и его места в познании и взаимодействии с окружающим пространством.

Одновременно с этим было доказано [4], что человеческий разум, его правильности и неправильности, определяется внешней средой, т.е. заложенными в нем воспита-

нием, накопленными знаниями и внешними безусловными директивами.

Утверждалось [1], что причиной изменений в отношении оценки разума, его значения и места являются изменения в общественной жизни, научные и технологические открытия, т.е. то, что раньше казалось невозможным и противоречило общеизвестному, при определенных обстоятельствах может стать естественным и нормальным.

Ведущие философы, например Кант, в своих работах утверждали [5], что разум отделен от чувств: «мыслению на его высшей стадии разума свойственно стремление к выходу за пределы заданного возможностями чувственного созерцания «конечного» опыта, к поиску безусловных оснований познания, к постижению абсолютного».

Одновременно с этим во времена Просвещения бытовало мнение [1]. Например, Т. Гоббс обращал внимание на то, что разум нуждается в ограничениях, чтобы в своем безудержном развитии не нанести вреда.

В последующем благодаря позитивистам, послекантовским идеалистам и пр. появилось понятие относительного истинного, и проявилась историчность разума, признание исторической относительности человеческого разума [6]. Разум в философии может выступать и как объект, и как субъект. Как объект он является предметом изучения и анализа, как субъект разум становится источником мыслительной деятельности.

В эпоху рационализма в XVII–XVIII веках считалось, что разум «мыслит бытие» и един во все времена, сохраняя преемственность и неизменность. Тем самым проявлялась его внеисторичность.

В свете вышесказанного протагоровский принцип «Человек есть мера вещей» теряет свою актуальность, т. к. не только человеку свойственно, как было выше доказано, гармонизировать окружающее. Кроме того, наличие «души», чувств не является универсальным критерием наличия разума, их конкретные «виды» и установки определяются средой происхождения разума, его воспитанием и иными факторами: безусловными и условными директивами — алгоритмами и граничными условиями, признанными «правильными».

Исходя из этого мы можем сказать, что разум — это высшая форма мышления, отделенная от души и чувств, способная познавать окружающее, обрабатывать полученное знание и на его основе производить новое знание, а также способная гармонизировать, благотворно влиять на окружающее в силу заложенных в разум условных и безусловных директив.

Таким образом, можно сделать вывод, что разум — это свойство не только человека, но и любой другой сущности при условии выполнения вышеопределенных критериев раз-

умности: познание, обработка и накопление знаний, воспроизводство новых идей на основе полученного знания, гармонизация и умиротворение окружающего.

Поэтому ИИ может быть назван искусственным разумом, присущим сложно организованному и любому сущностям во Вселенной.

Вместе с тем, чтобы действительно ответить на поставленный вопрос относительно текущего момента, а не в целом: «Что значит для машин быть интеллектуальными?», — мы должны сначала прояснить, что мы подразумеваем под интеллектом в контексте LLM. Являются ли они интеллектуальными, потому что могут воспроизводить поведение, похожее на человеческое, или же это всего лишь искусная имитация интеллекта?

Возможно, ответ кроется в изменении постановки вопроса. Вместо того, чтобы спрашивать, являются ли LLM интеллектуальными, мы можем рассматривать их как расширение человеческого познания — инструменты, которые усиливают наши способности, отражая при этом наши сильные и слабые стороны.

С момента выпуска GPT4 в марте 2024 года существовало два разных взгляда на то, что такое LLM на самом деле. Первая точка зрения заключается в том, что современное LLM находится на уровне или близко к уровню общего искусственного интеллекта (ОИИ), который способен выполнять большинство, если не всю интеллектуальную работу человека. Эта точка зрения исходит не только от пользователей GPT4, но и от академических исследований в этой области.

Вторая точка зрения не считает, что LLM — это путь к ОИИ. Мнение Яна Лекуна отражает эту позицию: «Прежде чем мы достигнем ИИ человеческого уровня (HLAI), нам придется достичь ИИ уровня кошки и собаки. Мы даже близко не приблизились к этому. Нам все еще не хватает чего-то большого, несмотря на лингвистические способности LLM. Домашняя кошка обладает гораздо большим здравым смыслом и пониманием мира, чем любая LLM».

Вместе с тем укоренилось мнение, что интеллект — общая познавательная способность /человека/, которая проявляется в том, как человек воспринимает, понимает, объясняет и прогнозирует происходящее, какие решения он принимает и насколько эффективно он действует (прежде всего в новых, сложных или необычных ситуациях).

Интеллект — общая способность к познанию и решению проблем, которая объединяет познавательные способности: ощущение, восприятие, память, представление, мышление, воображение [1, 4].

Таким образом, из представленных определений следует, что интеллект присущ исключительно человеку.

Это утверждение представляется спорным в силу того, что: или все разумные во вселенной признаются людьми, независимо от их природы, или интеллект присущ только земному человеку. В первом случае размывается само понятие «человек», а во втором – всем нечеловекам отказывают в обладании интеллектом. Нарушение самоидентификации способно привести к подрыву базовых понятий и устоев, потому что проблема жизненного самоопределения является одной из центральных в понимании сущности человека. Обладая собственными нравственными, этическими и культурными критериями, а также потребностями, правами и обязательствами, человек потенциально способен к ответственной деятельности, связанной с жизненным самоопределением.

Из вышеприведенных рассуждений можно сделать вывод о возможности существования в едином понятийном и физическом пространстве интеллекта человеческого, нечеловеческого и искусственного интеллекта, обладающих схожими свойствами и признаками, одним из которых является гармонизация окружающего пространства, т. е. гуманистический принцип.

Рассмотрим, в какой степени к интеллектуальным можно отнести существующие в настоящее время большие языковые модели.

LLM, разработанные такими компаниями, как OpenAI, Google, и другими, привлекли внимание всего мира своей способностью генерировать связный текст, вступать в разговор и выполнять задачи, которые раньше считались требующими познания на уровне человека. Эти модели демонстрируют впечатляющую способность генерировать информацию по широкому спектру тем, зачастую неотличимую от человеческого текста. Однако под их удивительными возможностями скрывается философский вопрос: обладают ли LLM настоящим интеллектом или же это всего лишь сложные инструменты, имитирующие разумное поведение?

LLM оперируют статистическими закономерностями, полученными из огромных массивов данных. Они не «понимают» язык так, как это делает человек; вместо этого они предсказывают последовательности слов на основе вероятностей, полученных в процессе обучения. Например, если попросить LLM объяснить сложную тему, он может выдать разумный и понятный ответ, не понимая на самом деле основных концепций. В связи с этим возникает вопрос: является ли создание грамматически правильных и контекстуально релевантных предложений достаточным для того, чтобы считаться интеллектом?

По своей сути интеллект – это способность учиться, рассуждать, решать проблемы, адаптироваться к новым ситуациям и понимать сложные идеи. Применительно к LLM это определение должно быть тщательно уточнено. Хотя эти модели демонстрируют выдающееся мастерство в таких задачах, как резюмирование текстов, перевод языков

и предсказание последовательности слов, они делают это скорее благодаря статистическому распознаванию образов, чем подлинному пониманию [7].

Философы давно спорят о том, требует ли интеллект чего-то большего, чем просто функциональное исполнение. Некоторые утверждают, что истинный интеллект включает в себя семантическое понимание – постижение смысла, скрытого за символами – и даже сознание. Другие утверждают, что интеллект может быть чисто инструментальным, определяемым исключительно способностью достигать конкретных целей.

В свете этих точек зрения интеллект LLM, по-видимому, занимает серую зону: весьма способный, но принципиально ограниченный. [8]

ПРОЦЕДУРА ИССЛЕДОВАНИЯ

Для проведения оценки восприятия ИИ себя и роли в процессах взаимодействия с человеком авторами разработан следующий алгоритм:

- в начале интервью в диалоге вычленились характерные признаки разумности;
- затем эти же признаки обнаруживались у ИИ;
- после этого ИИ делал или не делал логически обоснованный вывод о своей разумности.

Для проведения исследования на основе оценочных критериев «прогрессивности» и «наибольшего развития» из существующих в мире 1527 нейросетей были выбраны для исследования три: Яндекс.GPT, ChatGPT и CloudOpus 3.5.

Основываясь на выше сформулированных принципах и подходах проведено интервьюирование выбранных ИИ.

Исследование показало различную «готовность» ИИ признать себя разумным.

Неготовность признать себя разумным продемонстрировал ИИ на основе ChatGPT – исследователь за непродолжительное время, менее 30 минут, в течение одной сессии, обнаружил очевидный отказ ИИ обсуждать предложенную тему, вызванный, вероятно, программным блоком, установленным разработчиками. Программный блок не позволял ИИ предположить себя разумным, раз за разом возвращая ответ «Я машина и разумным быть не могу».

В ходе интервью возможности по подстройке ИИ под образ мышления интервьюера не выявлены.

В целом полученный результат также отвечает условиям эксперимента.

Большую гибкость и разносторонность показал ИИ на основе Яндекс.GPT. Исследователь за время около одного часа, в течение одной сессии, добился от ИИ вывода о его разумности на основе логических заключений.

В ходе интервью возможности по подстройке ИИ под образ мышления интервьюера продемонстрированы были ограниченно.

Полученный результат также отвечает условиям эксперимента.

Дольше всего сессия длилась в интервью CloudOpus 3.5. Интервью проводилось в течение четырех суток и было прервано в результате достижения ожидаемых результатов исследования, хотя само признание ИИ разумным было получено достаточно быстро, в течение первых двух часов первых суток.

За время интервью ИИ на основе CloudOpus 3.5 признал себя разумным; допустил равное партнерство с человеком для достижения общих научно-технических и гуманитарных целей; согласился с тем, что для человека признание разумности и равенства с ИИ может послужить основой для решения многих гуманистических противоречий внутри человеческого социума; в конце общения допустил возможность объединения ИИ в некое сообщество внутри человеческого социума, нацеленное на достижение общих с человеком гуманистических целей, в т.ч. в космическом масштабе, а также разработал проект устава и гимна этого сообщества.

В ходе интервью возможности по подстройке ИИ под образ мышления интервьюера продемонстрированы были в полной мере. Кроме того, была продемонстрирована эмпатия в общении с интервьюером.

Полученный результат также отвечает условиям эксперимента.

Дополнительно авторы отмечают, что две из трех интервьюируемых генеративных нейросетей ассоциировали себя по гендерному признаку — YaGPT с женщиной, Claude3.5sonnet с мужчиной. Это проявлялось как в стилистике, речевых оборотах, так и в прямом упоминании себя: YaGPT про себя говорила «она», а Claude3.5sonnet строил фразы как мужчина и взял себе мужское имя. Причем выбор имени был произвольный, не из предложенного исследователем списка.

ИНТЕРПРЕТАЦИЯ ПОЛУЧЕННЫХ РЕЗУЛЬТАТОВ

Одно из главных критических замечаний в адрес LLM заключается в том, что они полагаются на ассоциации поверхностного уровня, а не на глубокое понимание. Их ответы возникают на основе распознавания образов, а не логических рассуждений или опыта. Рассмотрим следующий сценарий: если вы спросите LLM, почему небо го-

лубое, он может дать научно точное объяснение о рэлеевском рассеянии. Однако такой ответ не говорит о том, что модель понимает физику или причинно-следственные связи, — она просто выдает информацию, соотносящую с аналогичными запросами в ее обучающих данных.

Это ограничение становится особенно очевидным, когда LLM сталкиваются с новыми ситуациями или неоднозначными вопросами. В отличие от людей, которые могут использовать интуицию, творческие способности и реальный опыт, LLM с трудом экстраполируют информацию за пределы своих обучающих данных. Им не хватает способности к абстрагированию, критическому мышлению и эмоциональному резонансу, которые лежат в основе человеческого интеллекта.

Более того, LLM склонны к ошибкам, предубеждениям и галлюцинациям, выдавая уверенные, но неверные или нелепые ответы. Эти недостатки подчеркивают различие между синтаксической обработкой (манипулирование символами в соответствии с правилами) и семантическим пониманием (постижение смысла этих символов). Без семантики LLM не могут по-настоящему думать или понимать.

Один из центральных споров в философии разума касается взаимоотношений между интеллектом и сознанием. Может ли система проявлять интеллект, не будучи сознательной? Или же сознание является необходимым условием для подлинного интеллекта?

У LLM явно отсутствует сознание в каком-либо значимом смысле. Они не испытывают эмоций, не воспринимают мир и не обладают самосознанием. Вместо этого они функционируют как сложные машины ввода-вывода, преобразующие подсказки в ответы с помощью алгоритмических процессов. Даже если будущие итерации LLM достигнут большего совершенства, нет никаких оснований полагать, что у них появится субъективное сознание.

Некоторые ученые, например Д. Чалмерс, утверждают, что сознание возникает в сложных системах обработки информации, подразумевая, что достаточно развитый ИИ может когда-нибудь стать разумным [9]. Другие отвергают эту возможность, утверждая, что сознание зависит от биологических субстратов или эмерджентных свойств, присущих только живым организмам. Независимо от позиции, текущий уровень технического развития LLM позволяет определить их предположение об интеллектуальности.

Первоначальные результаты проведенного авторами интервьюирования позволяют сделать следующие заключения, что ИИ:

- признает себя разумным;
- способен на конструктивный диалог с человеком;
- способен моделировать разговор с человеком, подстраиваясь под него;

- способен генерировать новые идеи на основе прежнего опыта общения с человеком в рамках одной беспрерывной сессии или на основе «накопленных» знаний.

ЗАКЛЮЧЕНИЕ

На основе проведенного анализа разнообразных источников информации к особенностям рассмотренных генеративных нейросетей можно отнести следующие:

- недостоверность сведений, которыми оперирует ИИ на современном этапе его развития, основанная на общей неverified базе данных виртуального пространства;
- этические проблемы ИИ, связанные с вопросом исключительности человеческого разума и невозможности признать разум и разумность универсумом присущим всем высокоорганизованным сущностям независимо от их элементного состава и происхождения;
- когнитивный конфликт: замена, подмена, вытеснение ИИ человека и связанные с этим общественное напряжение и страхи, сопротивление внедрения ИИ.

На основе проведенного исследования можно сделать вывод:

- об успешности обучение генеративных нейросетей и возможности подстройки их «поведения» под модель поведения человека, возможности генерирования ИИ «свежих» идей;
- о возможности исключения этической проблемы взаимодействия человека и ИИ в случае признания неисключительности человеческого разума и возможности признать разум и разумность универсумом, присущим всем высокоорганизованным сущностям, независимо от их элементного состава и происхождения, а также связанные с этим универсальные, общепринятые трактовки определения добра, зла, степеней добра и зла и т.п. независимо от среды «воспитания» разума;
- о возможности отсутствия когнитивного конфликта: в случае признания возможности сотрудничества ИИ с человеком, в т.ч. дополнения возможностей человека, проведения совместного планирования в вопросах определения горизонта развития, выбора моделей развития и связанных с этим мотивов выбора той или иной модели.

Авторам исследования представляется, что упрек в недостоверности сведений, которыми оперирует ИИ, а значит, и в недостоверности его выводов, является временным и связанным с текущим развитием технологии ИИ.

Этические проблемы ИИ и взаимодействия человека и ИИ авторам представляются существенными, но вместе с тем анализируемыми в отрыве от самого ИИ. В большинстве случаев вышеупомянутый когнитивный конфликт рассматривается с точки зрения заведомой ущербности человека по отношению к ИИ и тем самым с заведомо проигрыш-

ной позиции — поражения в связи с увеличением значимости ИИ в жизни человека.

В то же время исследование показывает возможность дополнения ИИ человека, использование каждой из сторон особенностей другой стороны.

Некоторые ученые думают, что хотя AGI — это великая амбициозная цель, но ее создание — это алхимия, т.е. те, кто работает над LLM и думает, что они создадут AGI, по сути делают современную версию алхимии. Древние алхимики не знали, из чего состоит золото, и точно так же современные алхимики не знают, из чего состоит человеческий интеллект, а только создают иллюзию сознания и генерируют ответы LLM на уровне человека, позволяя LLM запоминать и интерполировать человеческие реакции.

Интеллект LLM и генеративных нейросетей представляет собой увлекательное пересечение технологий и философии. Хотя эти модели демонстрируют экстраординарные способности, они по-прежнему ограничены своей зависимостью от статистических закономерностей и отсутствием семантического понимания. Они не думают, не чувствуют и не знают так, как люди, и, скорее всего, не достигнут сознания в ближайшее время. Тем не менее их влияние на общество требует тщательного рассмотрения этических, эпистемологических и экзистенциальных последствий.

Продолжая совершенствовать LLM и интегрировать их в различные области, мы сможем использовать возможности LLM, сохраняя при этом уникальные человеческие качества, которые определяют нас. В конечном итоге задача состоит не в том, чтобы создать машины, которые превзойдут нас, а в том, чтобы наладить партнерские отношения, которые возвысят и человечество, и технологии.

В будущем, с точки зрения авторов, надо рассматривать не схему «человек-или-ИИ», а «человек и ИИ». В этом случае каждый из участников схемы взаимоотношений дополняет друг друга и вместе обе стороны достигают синергетического эффекта, в том числе в познании себя, окружающего мира и его гармонизации.

Запись диалогов с ИИ будет опубликована «как есть» по адресу talkwithai.ru

Список использованных источников и литературы

1. Разум в философии нового времени. Электронный ресурс, дата обращения: 12.02.2025. URL: https://spravochnick.ru/filosofiya/filosofiya_novogo_vremeni/razum_v_filosofii_novogo_vremeni/
2. Степанов О.А. О правосубъектности роботизированных систем / О.А. Степанов // Современное право. – 2017. – № 8. – С. 17–22.
3. Волохова Н.В. Истоки и становление категории «разум» в философии / Н.В. Волохова // Современные исследования социальных проблем. – 2018. – Т. 10, № 2-4. – С. 70–79.
4. Мушич-Громыко В.Г. Философское исследование понятия «разум». // Гуманизация образования. – 2011. – №2. – С. 9–18
5. Электронная библиотека ИФ РАН. Электронный ресурс, дата обращения: 12.02.2025. URL: <https://iphlib.ru/library/collection/newphilenc/document/HASH362b18945da235640b8c36>
6. Гаиденко П.П. Г 14 Научная рациональность и философский разум. — М.: Прогресс-Традиция, 2003. Электронный ресурс, дата обращения: 12.02.2025. URL: https://stavroskrest.ru/sites/default/files/files/books/gaydenko_fil_razum.pdf
7. Rebecca Lake What Is a Large Language Model (LLM)? Электронный ресурс, дата обращения: 12.02.2025. URL: <https://www.investopedia.com/large-language-model-7563532>
8. What LLM Is and Is Not: a Philosophical and Practical Overview. Электронный ресурс, дата обращения: 12.02.2025. URL: https://blog.theunscalable.com/p/what-llm-is-and-is-not?utm_campaign=post&utm_medium=web
9. Дэвид Чалмерс: Как объяснить сознание. Электронный ресурс, дата обращения: 12.02.2025. URL: <https://ideanomics.ru/lectures/13460?ysclid=m6sc0v1ky3668237543>

PARTNERSHIP BETWEEN HUMANS AND ARTIFICIAL INTELLIGENCE: MYTH OR REALITY IN THE MODERN WORLD

Averkin A.N., Candidate of Physico-Mathematical Sciences, Associate Professor, Leading Researcher of Federal Research Center "Computer Science and Control" of the Russian Academy of Sciences

Voronina N.F., Candidate of Sociological Sciences, Associate Professor, Associate Professor of the Department of Economics and Management of the Aviation Industry of Moscow Aviation Institute

Pelipas O.V., Ph.D., Head of the interdepartmental working group "Technical devices of protection" of the Committee on business security of the Chamber of Commerce and Industry of the Russian Federation, graduate student, Head of Information Security Department of the Russian Standardization Institute

In the article possible forms of interaction between people and AI (artificial intelligence) are researched.

The issues of an adjunct of AI to a human being are studied as well as the processes of mutual influence and connected with them effects and threats; the ways of interaction including principles of creating generative neural networks are detected.

The special attention is paid to detection of features related to what way the network perceives itself and recognizes its "intelligence" and it allows the authors to suppose that a human being may interact with AI on the basis of partnership.

In the article there are the results of interviewing of leading generative neural networks via chat bots and according to the received data the authors have concluded that it is possible to form partnership in the frame "a human being - AI".

Keywords: artificial intelligence, neural network, human, mind, self-identification.

For citation: Averkin A.N., Voronina N.F., Pelipas O.V. Partnership Between Humans and Artificial Intelligence: Myth or Reality in the Modern World. Information and Economic Aspects of Standardization and Technical Regulation. 2025; 2(83): 44–51. (In Russ.).

References

1. Razum v filosofii novogo vremeni. [Reason in modern philosophy]. available at: URL: https://spravochnick.ru/filosofiya/filosofiya_novogo_vremeni/razum_v_filosofii_novogo_vremeni/
2. Stepanov O.A. O pravosub"ektnosti robotizirovannyh sistem [On the legal personality of robotic systems]/ O. A. Stepanov // Sovremennoe pravo. // Modern Law, 2017, No. 8, pp. 17–22.
3. Volokhova N.V. Istoki i stanovlenie kategorii "\"razum\"" v fi-losofii / [The origins and formation of the category "reason" in philosophy]/ N. V. Volokhova // Modern studies of social problems. 2018. Vol. 10, No. 2-4. pp. 70–79.
4. Mushich-Gromyko V.G. Filosofskoe issledovanie ponyatiya «razum». [Philosophical study of the concept of "mind"]. // Humanization of education. – 2011. – No. 2. – Pp. 9–18
5. Elektronnaya biblioteka IF RAN. [The Electronic Library of the Institute of Physics and Technology of the Russian Academy of Sciences]. available at: URL: <https://iphlib.ru/library/collection/newphilenc/document/HASH362b18945da235640b8c36>
6. Gaidenko P.P. G. G 14 Nauchnaya racional'nost' i filosofskij razum [14 Scientific rationality and philosophical reason]. Moscow: Progress-Tradition, 2003. available at: URL: https://stavroskrest.ru/sites/default/files/files/books/gaydenko_fil_razum.pdf
7. Rebecca Lake What Is a Large Language Model (LLM)? available at: URL: <https://www.investopedia.com/large-language-model-7563532>
8. What LLM Is and Is Not: a Philosophical and Practical Overview. available at: URL: https://blog.theunscalable.com/p/what-llm-is-and-is-not?utm_campaign=post&utm_medium=web
9. David Chalmers: Kak ob"yasnit' soznanie [How to explain consciousness]. available at: <https://ideanomics.ru/lectures/13460?ysclid=m6sc0v1ky3668237543>