Философия науки и техники
2026. Т. 31. № 1. С. 81–94
УДК 1(091)

Philosophy of Science and Technology
2026, vol. 31, no. 1, pp. 81–94
DOI: 10.21146/2413-9084-2026-31-1-81-94

ЭПИСТЕМОЛОГИЯ И КОГНИТИВНЫЕ НАУКИ

В.В. Робустова, А.А. Танюшина

«Инопланетный разум»: поиск новых метафор в восприятии искусственного интеллекта*

Робустова Вероника Валентиновна – кандидат филологических наук, доцент кафедры лингвистики, перевода и межкультурной коммуникации факультета иностранных языков и регионоведения. Московский государственный университет им. М.В. Ломоносова. Российская Федерация, 119991, г. Москва, Ленинские горы, д. 1; e-mail: nikarbs@yandex.ru

Танюшина Александра Александровна – кандидат философских наук, ассистент кафедры истории и теории мировой культуры философского факультета. Московский государственный университет им. М.В. Ломоносова. Российская Федерация, 119991, г. Москва, Ленинские горы, д. 1; e-mail: a.tanyushina@gmail.com

Статья посвящена анализу феномена антропоморфизма в контексте социокультурного восприятия искусственного интеллекта и проблеме культурологической интерпретации последнего. Несмотря на широкое распространение тенденции наделять ИИ человеческими качествами, такая практика сталкивается с принципиальными ограничениями, обусловленными радикальным отличием машинного интеллекта от человеческого. Современные ИИ-структуры, особенно децентрализованные мультиагентные модели и «рассуждающие» языковые системы, демонстрируют принципиально иную логику работы, основанную на статистической обработке данных и алгоритмических процессах, что делает их не подходящими для описания через антропоцентрические метафоры. В статье рассматриваются альтернативные подходы к пониманию природы ИИ, включая использование концепций «инопланетного интеллекта», экосистем и самоорганизующихся структур, а также исследуются параллели между ИИ и изучением роевого поведения у животных, демонстрирующего распределенную природу интеллекта. Подчеркивается не только необходимость преодоления чрезмерной антропоморфизации для более точного научного понимания и этически корректного использования ИИ, но также эпистемическая значимость

* Исследование выполнено при поддержке Междисциплинарной научно-образовательной школы Московского университета «Сохранение мирового культурно-исторического наследия» (при поддержке Программы развития МГУ, проект № 23-Ш02-05).

© Робустова В.В., 2026
© Танюшина А.А., 2026

сохранения антропоцентричных метафор при описании принципов работы искусственных систем для прогресса в когнитивных науках. Исследование предлагает переосмыслить культурные нарративы о машинном интеллекте через призму множественности форм разума, что способствует развитию новых теоретических и практических подходов к взаимодействию человека и ИИ в современном обществе.

Ключевые слова: искусственный интеллект, концептуальная теория метафоры, антропоморфизм, эпистемический мост, машинный разум, метафорический плюрализм, «инопланетный интеллект», «инструментальный интеллект»

Введение

Исследование эпистемологических оснований искусственного интеллекта в контексте его метафорического конструирования – актуальная междисциплинарная задача, в которой пересекаются интересы истории науки, лингвистики, когнитивистики и теории ИИ. Центральным объектом анализа в настоящей работе выступает метафора «искусственный интеллект» как концептуальный инструмент, опосредующий восприятие и интерпретацию феномена ИИ в научном и общественном дискурсах. Несмотря на широкое распространение данного термина, его семантическая неоднозначность и историческая изменчивость порождают ряд теоретических противоречий, требующих критического осмысления.

Во-первых, необходимо отметить, что рассматриваемая метафора изначально формировалась под влиянием антропоморфных и антропоцентрических установок, укорененных в западной интеллектуальной традиции. Уже в работах А. Тьюринга и Н. Винера прослеживается тенденция к проецированию на технические системы свойств человеческого разума, таких как мышление, обучение или принятие решений [Тьюринг, 1960; Wiener, 1948]. Подобная метафоризация, с одной стороны, способствовала популяризации идеи ИИ, делая его интуитивно понятным для широкой аудитории, с другой – создавала иллюзию онтологической эквивалентности между искусственными системами и человеческим сознанием. Однако по мере развития ИИ-технологий, особенно в эпоху глубокого обучения и нейроморфных вычислений, стала очевидной ограниченность данной метафоры: современные алгоритмы функционируют на принципах, радикально отличающихся от биологических механизмов мозга, что ставит под сомнение валидность антропоморфных аналогий.

Во-вторых, метафора «искусственный интеллект» тесно коррелирует с так называемой компьютерной метафорой, доминировавшей в когнитивной науке второй половины XX в. Эта концептуальная рамка, представляющая мозг как вычислительное устройство, а ментальные процессы – как алгоритмические операции, оказала значительное влияние на формирование исследовательских программ в области ИИ, однако в последние десятилетия подверглась критике со стороны энактивистских и воплощенных (embodied) подходов к сознанию, подчеркивающих роль телесности, сенсорики и социокультурного контекста в формировании когнитивных процессов. В результате возникает эпистемологический разрыв: ИИ, концептуализированный через устаревшую компьютерную метафору, оказывается неспособен адекватно моделировать ключевые аспекты

человеческого интеллекта, такие как эмоциональный интеллект, творчество или моральное суждение.

В-третьих, метафоры, используемые для описания ИИ, не являются нейтральными риторическими приемами, но обладают собственной эпистемической силой, формируя исследовательские программы, направляя финансирование и определяя общественные ожидания. Например, метафора «обучения» нейронных сетей имплицитно предполагает наличие у ИИ интенциональности и способности к самоорганизации, хотя в действительности речь идет о статистической оптимизации параметров модели. Подобные коннотации не только искажают публичное восприятие ИИ, порождая нереалистичные страхи или завышенные ожидания, но и влияют на саму научную практику, провоцируя исследователей на некритическое приписывание алгоритмам свойств, присущих исключительно человеческому мышлению.

Наконец, в-четвертых, актуальность критического анализа метафорического аппарата, используемого для описания ИИ, усиливается в условиях современного ИИ-бума, когда технологии машинного обучения проникают во все сферы общественной жизни – от медицины до искусства. Поскольку от выбора терминологии зависят этические и юридические рамки применения ИИ-систем, исследование приобретает не только теоретическое, но и практическое значение. Например, использование термина «автономный ИИ» может повлечь за собой ошибочную атрибуцию алгоритмам моральной ответственности, в то время как понятие «инструментальный ИИ» подчеркивает его подчиненную роль в процессах принятия решений.

Таким образом, исследование, с результатами которого мы собираемся познакомить читателей, имело своей целью:

1) рассмотреть теорию концептуальной метафоры и описать роль антропоморфных метафор в истории науки и философии техники;

2) проследить историческую эволюцию метафоры «искусственный интеллект» в ее взаимосвязи с доминирующими концепциями сознания;

3) выявить эпистемологические ограничения антропоморфного подхода к описанию ИИ;

4) рассмотреть существующие сегодня альтернативные метафорические фреймы, более адекватные современным ИИ-технологиям;

5) оценить положительные и отрицательные последствия возможной интеграции неантропоморфных метафор в современный научный и социокультурный дискурс.

Методологическая основа исследования – синтез историко-научного анализа, концептуального моделирования и критической теории метафоры – позволяет раскрыть динамику взаимодействия между научными репрезентациями ИИ и социокультурным контекстом их производства.

Метафорические траектории ИИ: от машин к нейросетям

Взаимодействие специалистов в области искусственного интеллекта и представителей лингвистической науки берет свое начало еще в 1950-х гг. Значительные достижения в области компьютерных наук последних десятилетий,

как уже было отмечено выше, способствовали широкому распространению и закреплению компьютерной метафоры мышления не только в массовом сознании, но и в профессиональном дискурсе когнитивной психологии.

Основополагающей работой в области изучения концептуальной метафоры является книга «Метафоры, которыми мы живем» Дж. Лакоффа и М. Джонсона. Предложенная авторами когнитивная теория метафоры строится на изучении процессов обработки структур знания: «Метафоризация основана на взаимодействии двух структур знаний – когнитивной структуры "источника" (source domain) и когнитивной структуры цели (target domain). В процессе метафоризации некоторые области цели структурируются по образцу источника, иначе говоря, происходит "метафорическая проекция" (metaphorical mapping) или "когнитивное отображение" (cognitive mapping)» [Лакофф, Джонсон, 2008, с. 9]. Исследователи трактуют область источника как более конкретное знание, которое человек получает в процессе познания действительности и непосредственного опыта взаимодействия с ней. Сфера цели видится как менее конкретное, менее определенное знание, для осмысления которого и привлекается концептуальная метафора. Таким образом, человеческий опыт базируется на структурных корреляциях, выступающих в качестве мотивационной основы метафорических переносов. Если такая мотивация детерминирована эмпирическим опытом индивида, метафора приобретает характер естественного феномена: «Данное явление объясняется тем, что корреляция между областью-источником и областью-целью мотивирована опытом, равно как и специфические характеристики рассматриваемого метафорического отображения» [Лакофф, 2004, с. 362].

Основные положения теории концептуальной метафоры, разработанной в трудах Лакоффа и Джонсона, могут быть систематизированы следующим образом:

– концептуальная система человека характеризуется имманентной метафоричностью, концептуальные метафоры выполняют структурирующую функцию, обеспечивают межконцептуальные связи и способствуют осмыслению абстрактных понятий;

– концептуальные метафоры образуют системные взаимосвязи, формируя кластеры вокруг культурно значимых концептов конкретной лингвокультурной общности, метафорические концепты репрезентируют культурно-специфическую картину мира, инкорпорируя ценностные ориентиры и аксиологические установки соответствующей культуры;

– метафоры коррелируют с совокупностью вербализующих их языковых единиц, что обеспечивает возможность исследования процессов метафоризации посредством анализа метафорических концептов, репрезентированных языковыми метафорами;

– метафорический перенос подчиняется принципу инвариантности (invariance principle), сохраняя когнитивную топологию области-источника;

– метафоры обладают способностью фокусировать внимание реципиента на определенных аспектах описываемого феномена, одновременно элиминируя другие его характеристики;

– генезис концептуальных метафор в когнитивной системе индивида обусловлен накоплением эмпирического опыта взаимодействия с объектами

действительности, метафоры выполняют структурирующую функцию в отношении мышления, поведенческих паттернов и деятельности человека.

Становление и последующая трансформация метафоры «искусственный интеллект» представляют собой сложный процесс, отражающий динамику философских, научных и техкультурных представлений о природе разума и возможностях его машинной репродукции. Историко-эпистемологический анализ позволяет выделить несколько ключевых этапов, на которых концептуализация ИИ оказывалась тесно сопряжена с доминирующими в соответствующий период моделями сознания, формируя своеобразную эпистемическую связку, определяющую как исследовательские стратегии, так и публичный дискурс.

Первоначальные попытки осмысления возможности создания «искусственного разума» восходят к механистической философии Нового времени, в рамках которой сознание интерпретировалось как продукт сложных, но принципиально познаваемых механических взаимодействий [Барышников, 2002]. Эта установка, восходящая к работам Р. Декарта и Ж.О. де Ламетри, нашла свое отражение в создании механических автоматов XVIII–XIX вв. (например, «Пишущий мальчик» П. Жаке-Дро), которые, имитируя человеческие действия, служили наглядной метафорой редукции интеллекта к физическим процессам [Шилов, 2019]. Однако вплоть до середины XX в. подобные аналогии оставались скорее философскими спекуляциями, нежели основой для конкретных технических проектов.

Систематическая концептуализация ИИ как научной и инженерной задачи началась в 1950-х гг. в контексте кибернетического движения, возглавляемого Н. Винером. Кибернетическая парадигма, рассматривающая живые организмы и машины как системы управления и обработки информации, основанные на принципах обратной связи, задала новую метафорическую рамку для осмысления ИИ. Интеллект интерпретировался как способность к адаптивному поведению в среде, достигаемому через алгоритмическую обработку входных сигналов и генерацию соответствующих выходных реакций. Показательно, что предложенный А. Тьюрингом в 1950 г. тест на определение «машинного интеллекта» был основан именно на этой бихевиористской установке, редуцирующей разум к наблюдаемому поведению.

Следующий этап в эволюции метафоры «искусственный интеллект» связан с развитием компьютерных наук в 1960–1970-х гг. и формированием так называемого символьного подхода. В этот период доминирующей моделью сознания становится уже упоминавшаяся выше компьютерная метафора. Эта концепция нашла отражение в работах А. Ньюэлла и Г. Саймона и привела к созданию экспертных систем и программ, основанных на формальной логике, которые рассматривались как первые практические реализации искусственного интеллекта. Однако уже к концу 1980-х гг. стали очевидны ограничения символьного подхода, неспособного моделировать такие аспекты человеческого интеллекта, как контекстуальная гибкость, интуиция или эмоциональная окрашенность мышления [Milkomski, 2018].

Кризис символьной парадигмы совпал с подъемом коннекционизма и развитием искусственных нейронных сетей, вдохновленных структурой биологического мозга. Эта смена метафорического фрейма – от «компьютера»

к «нейронной сети» – ознаменовала переход от дискретных, логически прозрачных моделей ИИ к распределенным, субсимволическим системам, способным к обучению на основе данных. Коннекционистская модель, опирающаяся на принципы параллельной обработки информации и адаптивной модификации связей между элементами, оказалась значительно ближе к современным нейробиологическим представлениям о работе мозга, чем предшествующие подходы. Однако и эта метафора не лишена противоречий: несмотря на внешнее сходство с биологическими нейронами, искусственные нейронные сети функционируют на основе математических абстракций, далеких от реальных биохимических и электрофизиологических процессов в мозге [Milkomski, 2018; Boden, 2008].

Современный этап эволюции метафоры ИИ характеризуется двумя противоположными тенденциями. С одной стороны, развитие глубокого обучения и генеративных моделей (например, моделей-трансформеров) привело к появлению антропоморфных терминов («галлюцинации», «внимание»), создающих иллюзию психологической эквивалентности между ИИ и человеком. С другой стороны, в философии сознания и когнитивной науке наравне с теориями вычислительного функционализма набирают силу энактивистские и процессуальные подходы, рассматривающие интеллект как результат динамического взаимодействия организма со средой, не сводимого к вычислительным процессам [Gallagher, 2023]. В этом контексте метафора «искусственный интеллект» все чаще заменяется более нейтральными терминами, такими как «машинное обучение» или «алгоритмические системы», что отражает стремление к деантропоморфизации ИИ и признанию его онтологической инаковости по отношению к человеческому разуму. Однако подобного рода метафоры также не очень удачны, т.к. имплицитно предполагают наличие у ИИ квазипсихологических свойств – интенциональности, субъективности или даже эмоциональности. Современные ИИ-системы, даже самые продвинутые, функционируют на основе статистической обработки данных и оптимизации математических функций, не обладая при этом ни сознанием, ни самосознанием, ни способностью к рефлексии, на что в том числе обращают внимание ведущие философы. Например, так называемые галлюцинации больших языковых моделей (LLM), часто интерпретируемые как «творческие ошибки», в действительности представляют собой результат вероятностной генерации текста, не имеющий ничего общего с человеческими воображением и заблуждением.

Второе ограничение касается эпистемологической редукции сложных технических процессов к упрощенным антропоморфным нарративам. Например, термин «глубокое обучение» метафорически отсылает к идее «глубины» человеческого понимания, однако в реальности «глубина» нейронной сети определяется исключительно количеством ее скрытых слоев, не имея ничего общего с семантической или концептуальной глубиной. Аналогично, понятие «внимание» (attention) в трансформерных архитектурах хотя и вдохновлено психологическим феноменом фокусировки сознания, в действительности представляет собой математическую операцию взвешивания входных данных. Эта редукция сложных когнитивных процессов к вычислительным процедурам

не только создает ложное впечатление о наличии у ИИ сознания или интенциональности, но и упрощает реальную природу человеческого внимания.

Наконец, стоит отметить что большинство терминов, используемых для описания ИИ, укоренены в западной интеллектуальной традиции, что создает барьеры для межкультурной коммуникации и ограничивает возможности альтернативных концептуализаций. Например, в японской культуре, где граница между человеком и машиной исторически менее жесткая (что отражено в концепциях «каракури» или «тама»[1]), метафора «искусственный интеллект» может приобретать иные коннотации, что ставит под вопрос универсальность антропоморфного подхода.

Таким образом, несмотря на свою историческую роль в популяризации области ИИ-исследований, антропоморфный подход к описанию ИИ сегодня превращается в эпистемологический барьер, препятствующий адекватному пониманию природы современных технологий машинного обучения. Преодоление этого барьера требует критического переосмысления используемого метафорического аппарата и разработки новых концептуальных рамок, способных отразить уникальные онтологические и функциональные характеристики искусственного интеллекта, не сводя их к антропоцентрическим проекциям.

За пределами антропоморфизма: новые метафоры для описания ИИ

Как уже было отмечено выше, сегодня в научном и философском дискурсе активно разрабатываются альтернативные метафорические фреймы, ориентированные на преодоление ограничений антропоцентризма и более адекватное отражение онтологической и функциональной специфики современных ИИ-систем. Эти концептуальные модели, варьирующиеся от радикально иных онтологических допущений до переосмысления самой идеи интеллекта, открывают новые перспективы для теоретического осмысления ИИ и его интеграции в социотехнические системы.

---

[1] Каракури – историко-культурный термин, обозначающий механические устройства, и прежде всего традиционные японские автоматы-куклы раннего Нового времени (период Эдо), предназначенные для зрелищной демонстрации «самодвижения» и иллюзии агентности. Каракури не сводится к «технике» в узком смысле. В культурной семантике это также режим восприятия механизма как носителя внутренней упорядоченности и квазиживого поведения, где важна не функциональная утилитарность, а способность артефакта производить эффект «оживления» и социально распознаваемой целенаправленности. «Тама» – многозначное понятие японской религиозно-философской лексики, связанное с представлениями о душе/духовной сущности, жизненной силе и агентности, прежде всего в синтоистских (а также в ряде буддийских и народных) контекстах. В классических и фольклорных моделях «тама» может мыслиться как не полностью тождественная индивидуальному Я субстанция, способная отделяться, перемещаться, усиливаться или ослабевать; она участвует в объяснении того, как предметы, места или существа оказываются «наделенными силой» и вызывают нормативные реакции уважения, осторожности либо ритуального обращения. В более широком культурном горизонте «тама» сопряжено с представлениями о возможности духовной «насыщенности» мира (близко к логике анимизма/панпсихизма в сравнительной религиоведческой перспективе), что концептуально смягчает жесткую дихотомию «одушевленное–неодушевленное».

Одной из наиболее популярных альтернатив является метафора «инопланетный интеллект» (alien intelligence), предложенная рядом исследователей для описания ИИ как принципиально чуждой человеческому разуму формы когнитивности [Sipper, 2018, web]. Эта метафора акцентирует внимание на том, что ИИ-системы, основанные на статистической обработке данных и оптимизационных алгоритмах, функционируют в соответствии с логикой, радикально отличной от человеческой рациональности, укорененной в биологической эволюции и культурном опыте. Концепция «инопланетного интеллекта» позволяет избежать проекции антропоморфных категорий (таких как «цели», «намерения» или «понимание») на ИИ, подчеркивая его инаковость и несводимость к человеческим когнитивным паттернам. В рамках этого подхода, например, способность больших языковых моделей генерировать связный текст интерпретируется не как «понимание» языка, а как результат высокомерной интерполяции в многомерном пространстве вероятностных распределений – процесс, не имеющий аналогов в человеческой психике.

Другой продуктивной метафорой является концепция «роевого интеллекта» (swarm intelligence), заимствованная из биологии и теории сложных систем. Эта модель, применяемая для описания децентрализованных ИИ-систем (например, алгоритмов оптимизации на основе колоний агентов или роевых робототехнических систем), акцентирует внимание на эмерджентных свойствах, возникающих в результате взаимодействия множества простых агентов [Rosenberg, 2016]. В отличие от антропоморфной метафоры «искусственный разум», подразумевающей централизованный контроль и иерархическую организацию, метафора «роевой интеллект» подчеркивает распределенный и самоорганизующийся характер ИИ-процессов. Это позволяет более адекватно описывать современные мультиагентные системы, где «интеллект» не локализован в отдельном элементе, а возникает как системный эффект коллективного поведения.

Смежной перспективной концептуализацией представляется метафора «эмерджентный (прорастающий) интеллект» (emergent intelligence), акцентирующая внимание на процессе постепенного формирования сложных когнитивных структур из простых взаимодействий [Pegluci, Boundry, 2011]. Эта концепция, вдохновленная теориями самоорганизации и автопоэзиса, рассматривает ИИ как динамическую систему, чья «интеллектуальность» не задана изначально, а возникает в процессе взаимодействия с окружающей средой и адаптации к ней. В рамках этого подхода, например, «обучение» нейронных сетей интерпретируется не как «запоминание» или «тренировка», а как процесс структурной сопряженности (structural coupling) системы с потоком данных, приводящий к эмерджентному возникновению новых функциональных свойств. Такая метафора позволяет избежать телеологических коннотаций, присущих антропоморфным описаниям, и более адекватно отражает нелинейный и непредсказуемый характер развития современных ИИ-систем.

Четвертой значимой альтернативой является метафора «инструментальный интеллект», предложенная рядом исследователей для описания ИИ как принципиально неантропоморфного феномена. Эта концепция рассматривает ИИ-системы не как «разумных агентов», а как сложные инструменты, расширяющие человеческие когнитивные и практические возможности. В частности,

системы поддержки принятия решений на основе машинного обучения интерпретируются не как «искусственные эксперты», а как высокоспециализированные инструменты анализа данных, чья эффективность зависит от контекста применения и качества исходной информации. Метафора «инструментальный интеллект» позволяет избежать этических и юридических коллизий, связанных с атрибуцией ИИ моральной ответственности, и фокусирует внимание на прагматических аспектах взаимодействия человека и машины.

Наконец, следует упомянуть концепцию «экологического интеллекта» (ecological AI), развивающуюся в рамках постгуманистической философии и теории техноценозов. Она рассматривает ИИ не как изолированную систему, а как элемент сложной социотехнической экосистемы, взаимодействующий с человеческими и нечеловеческими акторами. Это позволяет, например, анализировать алгоритмы рекомендательных систем в качестве не «манипуляторов» общественным мнением, а компонентов динамической информационной экосистемы, формирующих и самих формируемых коллективными паттернами поведения пользователей. Такая перспектива предполагает более адекватное описание эффектов коэволюции ИИ и общества, свободное от упрощенных дихотомий типа «человек vs машина».

Каждая из упомянутых альтернативных метафор обладает своими сильными сторонами и ограничениями, однако их совокупное применение создает многомерное концептуальное пространство, благодаря которому можно более точно картографировать ландшафт современных ИИ-технологий. Например, комбинация метафор «инопланетный интеллект» и «роевой интеллект» может быть продуктивна для описания гибридных человеко-машинных систем, где ИИ-компоненты функционируют по логике, чуждой человеческой, но при этом интегрированы в распределенные сети принятия решений. Аналогично синтез концепций «прорастающего» и «экологического» интеллекта позволяет моделировать долгосрочные процессы интеграции ИИ в социальные и экономические структуры, учитывая эффекты эмерджентности и коэволюции.

Важно отметить, что разработка альтернативных метафорических фреймов не является чисто академическим упражнением, но имеет непосредственные практические импликации. Например, переход от антропоморфной метафоры «автономный ИИ» к концепции «инструментального интеллекта» может радикально изменить подходы к регуляции ИИ-технологий, сместив фокус с абстрактных дискуссий о «правах роботов» на конкретные механизмы обеспечения надежности и прозрачности алгоритмических систем. Аналогично внедрение метафоры «экологический интеллект» в практику разработки ИИ может способствовать созданию более устойчивых и адаптивных социотехнических систем, учитывающих сложные петли обратной связи между технологиями и обществом.

Особенности метафорического сдвига в ИИ-дискурсе

Интеграция неантропоморфных метафорических фреймов в научный и социокультурный дискурс об искусственном интеллекте – амбивалентный процесс, порождающий как значительные эпистемологические и практические

преимущества, так и комплекс проблем, требующих критического осмысления. С одной стороны, преодоление антропоцентрических установок открывает новые горизонты для теоретического моделирования ИИ и его интеграции в сложные социотехнические системы. С другой стороны, радикальный разрыв с привычными метафорами может привести к коммуникационным барьерам и этическим коллизиям, особенно в областях, где антропоморфные интерпретации ИИ уже глубоко укоренились в общественном сознании.

Одним из ключевых позитивных последствий является формирование эпистемического моста между антропоморфным пониманием ИИ и вычислительными моделями работы человеческого мозга. Неантропоморфные метафоры, такие как «инопланетный интеллект» или «роевая когнитивность», создают концептуальное пространство для сравнительного анализа биологических и искусственных когнитивных систем без редукции одних к другим. Например, концепция «прорастающего интеллекта» позволяет интерпретировать обучение нейронных сетей как процесс, аналогичный формированию нейронных связей в развивающемся мозге, но при этом подчеркивает принципиальную инаковость механизмов, лежащих в основе этих процессов. Такой подход способствует преодолению дихотомии «естественное – искусственное» и создает предпосылки для разработки гибридных моделей интеллекта, сочетающих преимущества биологических и вычислительных систем (биокомпьютеров).

В практическом плане неантропоморфные метафоры могут радикально трансформировать подходы к проектированию и регуляции ИИ-систем. Например, переход от концепции «автономного ИИ» к модели «инструментального интеллекта» смещает фокус разработчиков с создания «самодостаточных агентов» на разработку контекстно-зависимых инструментов, чья функциональность жестко привязана к конкретным задачам и условиям применения. Это не только повышает надежность и прозрачность ИИ-систем, но и снижает риски неконтролируемого поведения, часто ассоциируемые с антропоморфными представлениями об автономности. Аналогично внедрение метафоры «экологический интеллект» в практику управления большими данными может способствовать созданию более устойчивых информационных экосистем, учитывающих сложные взаимодействия между алгоритмами, пользователями и социальными институтами.

Однако интеграция неантропоморфных концептов сопряжена и с рядом проблем. Во-первых, радикальный разрыв с антропоморфными метафорами может привести к коммуникационным барьерам между экспертным сообществом и широкой общественностью. Например, замена в публичном дискурсе термина «искусственный интеллект» на более точный с технической точки зрения термин «вычислительные оптимизационные системы» рискует сделать обсуждение ИИ-технологий менее доступным для неспециалистов, что может замедлить формирование общественной рефлексии по поводу их этических и социальных импликаций. Во-вторых, неантропоморфные метафоры, особенно радикальные варианты вроде «инопланетного интеллекта», могут спровоцировать новые формы технофобии, основанные на восприятии ИИ как принципиально чуждой и потенциально враждебной сущности.

Особого внимания заслуживает амбивалентная роль антропоморфных метафор в сферах, где они уже глубоко интегрированы в социальные практики – маркетинге, бизнесе и популярной культуре. С одной стороны, антропоморфизация ИИ выполняет здесь важные прагматические функции: делает сложные технологии более понятными для потребителей, создает эмоциональную связь с брендами (например, голосовые помощники с «личностью») и стимулирует креативность в индустрии развлечений (от фильмов о «мыслящих роботах» до ИИ-персонажей в видеоиграх). С другой стороны, она формирует нереалистичные ожидания относительно возможностей ИИ, что может приводить к разочарованиям (например, в случае неадекватного поведения чат-ботов) или даже к манипуляциям общественным мнением через создание иллюзии объективности алгоритмических решений.

В этом контексте особую актуальность приобретает задача разработки гибких метафорических стратегий, способных сочетать эвристический потенциал неантропоморфных концептов с прагматическими преимуществами антропоморфных нарративов. Одно из перспективных направлений – создание многоуровневых описаний ИИ, где неантропоморфные метафоры используются в экспертном дискурсе для теоретического моделирования и технической документации, в то время как антропоморфные интерпретации сохраняются в публичной коммуникации, но с обязательной экспликацией их метафорического статуса.

Другим важным направлением является развитие метафорического плюрализма – одновременное использование нескольких концептуальных рамок для описания одних и тех же ИИ-систем в зависимости от контекста. Например, алгоритмы рекомендательных систем могут интерпретироваться как «роевая когнитивность» при анализе их коллективного поведения, как «инструментальный интеллект» при оценке их функциональности и как элемент «экологического интеллекта» при изучении их долгосрочного воздействия на информационную среду. Такой подход не только обогащает аналитический инструментарий исследователей, но и способствует формированию более нюансированного общественного понимания ИИ как многогранного феномена, не сводимого к однозначным определениям.

Заключение

Историческая траектория метафоры «искусственный интеллект» демонстрирует ее глубокую укорененность в доминирующих эпистемологических парадигмах своего времени. От механистических автоматов XVIII в. до нейроморфных чипов XXI в. каждая новая технологическая и концептуальная модель сознания порождала соответствующую метафору ИИ, которая, в свою очередь, становилась инструментом для переосмысления самой природы человеческого разума. Этот диалектический процесс, в котором технические инновации и философские теории взаимно конституируют друг друга, продолжает определять развитие ИИ-дискурса, ставя перед исследователями задачу критического анализа эпистемологических оснований используемых метафор и их импликаций для будущего человеческой цивилизации.

Стоит отметить, что поиск альтернативных метафорических оснований для концептуализации ИИ представляет собой не просто лингвистическое упражнение, но фундаментальный эпистемологический проект, направленный на преодоление антропоцентрических ограничений и создание более адекватного языка для описания нечеловеческих форм интеллекта. Этот процесс, требующий междисциплинарного синтеза лингвистики, когнитивной науки, философии и инженерии, не только обогащает теоретические основания исследовательских программ, но и создает предпосылки для более ответственной и этически обоснованной интеграции искусственного интеллекта в глобальную социотехническую систему.

Распространение неантропоморфных метафор в дискурсе об ИИ – сложный диалектический процесс, требующий баланса между эпистемологической строгостью и коммуникативной эффективностью, между теоретической инновационностью и практической применимостью. Успех этого процесса будет во многом зависеть от способности научного сообщества, технологических компаний и общественных институтов выработать гибкие стратегии метафорического моделирования, способные адаптироваться к быстро меняющемуся ландшафту ИИ-технологий, сохраняя при этом философскую рефлексию и критическую дистанцию по отношению к собственным концептуальным построениям. В конечном счете речь идет не о замене одних метафор другими, а о создании динамической системы координат, позволяющей ориентироваться в пространстве нечеловеческих форматов рациональности, не теряя связи ни с технической реальностью ИИ-систем, ни с фундаментальными ценностями человеческой цивилизации.

Список литературы

Барышников, 2002 – *Барышников П.Н.* Разум как машина: влияние механицизма на концептуальные основания компьютерной метафоры // Вестник РУДН. Сер. Философия. 2002. Т. 26. № 4. С. 755–769.

Лакофф, 2004 – *Лакофф Д.* Женщины, огонь и опасные вещи: Что категории языка говорят нам о мышлении / Пер. с англ. И.Б. Шатуновского. М.: Языки славянской культуры, 2004. 792 с.

Лакофф, Джонсон, 2008 – *Лакофф Д., Джонсон М.* Метафоры, которыми мы живем / Пер. с англ. А.Н. Баранова и А.В. Морозовой. Под ред. и с предисл. А.Н. Баранова. 2-е изд. М.: Издательство ЛКИ, 2008. 256 с.

Тьюринг, 1960 – *Тьюринг А.М.* Может ли машина мыслить? / Пер. с англ. Ю.А. Данилова. М.: Государственное издательство физико-математической литературы, 1960. 67 с.

Шилов, 2019 – *Шилов В.* На пути к искусственному интеллекту: Логические машины и их создатели. М.: ЛЕНАНД, 2019. 244 с.

Boden, 2008 – *Boden M.A.* Mind as Machine. A History of Cognitive Science. Oxford: Clarendon Press, 2008. 1631 p.

Milkomski, 2018 – *Milkomski M.* From Computer Metaphor to Computational Modeling: The Evolution pof Computationalism // Minds and Machines. 2018. No. 3. P. 515–541.

Pigluici, Boudry, 2011 – *Pigluici M., Boudry M.* Why Machine-Information Metaphors are Bad for Science and Science Education // Science & Education. 2011. No. 20 (5–6). P. 471–484.

Gallagher, 2023 – *Gallagher S.* Embodied and Enactive Approaches to Cognition. Cambridge: Cambridge University Press. 2023. 88 p.

Rosenberg, 2016 – *Rosenberg L.* Artificial Swarm Intelligence, a Human-in-the-Loop Approach to A.I. // Proceedings of the AAAI Conference on Artificial Intelligence. 2016. Vol. 30. No. 1. P. 4381–4382.

Sipper, 2018, web – *Sipper M.* "A" is for Alien: Analogies Between Modern AI and Richard Feynman's "The Character of Physical Law" / Readmedium. 2018. URL: https://readmedium.com/a-is-for-alien-analogies-between-modern-ai-and-richard-feynmans-the-character-of-physical-law-57b5cbe93cf8 (дата обращения: 18.01.2026).

Wiener, 1948 – *Wiener N.* Cybernetics: Or Control and Communication in the Animal and the Machine. P.: Hermann & Cie; Cambridge, Mass.: MIT Press, 1948. 352 p.

"Alien intelligence": searching for new metaphors in the perception of artificial intelligence

Veronika V. Robustova

Lomonosov Moscow State University. GSP-1, Leninskie Gory, Moscow, 119991, Russian Federation; e-mail:nikarbs@yandex.ru

Alexandra A. Tanyushina

Lomonosov Moscow State University. GSP-1, Leninskie Gory, Moscow, 119991, Russian Federation; e-mail: a.tanyushina@ gmail.com

This article analyzes the phenomenon of anthropomorphism in the context of sociocultural perception of artificial intelligence and the problem of its culturological interpretation. Despite the widespread tendency to attribute human qualities to AI, this practice encounters fundamental limitations due to the radical differences between machine and human intelligence. Contemporary AI structures, particularly decentralized multi-agent models and "reasoning" language systems, demonstrate fundamentally different operational logic based on statistical data processing and algorithmic procedures, making them unsuitable for description through anthropocentric metaphors. The article examines alternative approaches to understanding the nature of AI, including the use of concepts such as "alien intelligence", ecosystems, and self-organizing structures, while also exploring parallels between AI and the study of swarm behavior in animals that demonstrate the distributed nature of intelligence. The study emphasizes not only the necessity of overcoming excessive anthropomorphization for more accurate scientific understanding of AI and ethically appropriate use of this technology, but also the epistemic significance of preserving anthropocentric metaphors for describing the operational principles of artificial systems for progress in cognitive sciences. The research proposes reconceptualizing cultural narratives about machine intelligence through the lens of multiple forms of mind, which contributes to the development of new theoretical and practical approaches to human-AI interaction in contemporary society.

Keywords: artificial intelligence, conceptual metaphor theory, anthropomorphism, epistemic bridge, machine mind, metaphorical pluralism, "alien intelligence", "instrumental intelligence"

Acknowledgements: This research has been supported by the Interdisciplinary Scientific and Educational School of Moscow University "Preservation of the World Cultural and Historical Heritage" (supported by the MSU Development Program, project No. 23-Ш02-05).

References

Baryshnikov, P.N. “Razum kak mashina: vliyanie mekhanicizma na konceptual’nye osnovaniya komp’yuternoj metafory” [Mind as a Machine: The Influence of Mechanism on the Conceptual Foundations of the Computer Metaphor], *Vestnik RUDN. Seriya: Filosofiya*, 2002, vol. 26, no. 4, pp. 755–769. (In Russian)

Boden, M.A. *Mind as Machine. A History of Cognitive Science*. Oxford: Clarendon Press, 2008. 1631 pp.

Lakoff, D. *Zhenshchiny, ogon’ i opasnye veshchi: Chto kategorii yazyka govoryat nam o myshlenii* [Women, Fire, and Dangerous Things: What Categories Reveal About the Mind], trans. by I.B. Shatunovskij. Moscow: Yazyki slavyanskoj kul’tury Publ., 2004. 792 pp. (In Russian)

Lakoff, D., Dzhonson, M. *Metafory, kotorymi my zhivem* [Metaphors We Live By], trans. by A.N. Baranov, A.V. Morozova. Second edition. Moscow: Izdatel’stvo LKI Publ., 2008. 256 pp. (In Russian)

Milkomski, M. “From Computer Metaphor to Computational Modeling: The Evolution of Computationalism”, *Minds and Machines*, 2018, no. 3, pp. 515–541.

Pigluici, M., Boudry, M. “Why Machine-Information Metaphors are Bad for Science and Science Education”, *Science & Education*, 2011, no. 20 (5–6), pp. 471–484.

Rosenberg, L. “Artificial Swarm Intelligence, a Human-in-the-Loop Approach to A.I.”, *Proceedings of the AAAI Conference on Artificial Intelligence*, 2016, vol. 30, no. 1, pp. 4381–4382.

Shilov, V. *Na puti k iskusstvennomu intellektu: Logicheskie mashiny i ih sozdateli* [Towards to Artificial Intelligence: Logical Machines and their Creators]. Moscow: LENAND Publ., 2019. 266 pp. (In Russian)

Sipper, M. “‘A’ is for Alien: Analogies Between Modern AI and Richard Feynman’s ‘The Character of Physical Law’”, *Readmedium*, 2018. URL: https://readmedium.com/a-is-for-alien-analogies-between-modern-ai-and-richard-feynmans-the-character-of-physical-law-57b5cbe93cf8 (accessed on: 18.01.2026).

Turing, A.M. *Mozhet li mashina myslit’?* [Can machines think?], trans. by Yu.A. Danilov. Moscow: Gosudarstvennoe izdatel’stvo fiziko-matematicheskoj literatury Publ., 1960. 67 pp. (In Russian)

Wiener, N. *Cybernetics: Or Control and Communication in the Animal and the Machine*. Paris: Hermann & Cie; Cambridge, Mass.: MIT Press, 1948. 352 pp.