

Перспективы развития ИИ: философский анализ гипотезы нахождения сознания вне тела человека.

Научный руководитель – Арапов Александр Владиленович

Алфимов Артём Александрович

Студент (бакалавр)

Воронежский государственный университет, Воронеж, Россия

E-mail: alfimov.aa20@gmail.com

В последнее десятилетие развитие искусственного интеллекта существенно изменило подход к пониманию фундаментальных проблем человека: проблемы познания и человеческого мышления. Современные системы ИИ демонстрируют сложные формы принятия решений и обработки информации. Именно из-за этого возникает вопрос: обладают ли нейросети даже малейшей частицей сознания в человеческом понимании, или всего лишь имитируют когнитивные процессы. Центральной проблемой становится различие между интеллектуальной деятельностью и субъективным опытом. Актуализируется в этой теме и вопрос об интеллекте и телесности как о достаточном условии возникновения сознания. Что именно необходимо для самосознания? Только ли взаимодействие с внешней средой влияет на развитие когнитивных процессов? В настоящее время различается три типа ИИ: Узкий (слабый) ИИ (англ. – Artificial Narrow Intelligence, ANI); Общий ИИ (англ. – Artificial General Intelligence, AGI); Суперсильный ИИ (англ. – Artificial Super Intelligence, ASI) [2, 37]. Узкий используется в аграрном секторе экономики, сборке и применении беспилотных летальных аппаратов, медицине и тд. Общий ИИ это известная всем нам языковая модель «Chat GPT», которая обучается на данных из общедоступного интернета. Суперсильный же ИИ можно объяснить на примере мысленного эксперимента «Василиск Рокко». Человечество в будущем будет способно создать такой искусственный интеллект, который сможет своим поведением подражать индивиду, в частности правителю-диктатору и «наказывать» тех, кто не поддерживал его создание. В научной работе «Когнитивизм как база искусственного интеллекта» [Там же, 39] авторы приходят к мысли о том, что для создания такого сверхсильного ИИ возможно человеку придется отойти от изученных концепций о работе мозга и реализовать то, что до этого считалось невозможным. Именно на этом этапе, по нашему мнению, происходит деление на несколько видов теорий о происхождении сознания. Даже если мы проведем мысленный эксперимент и представим, что сознание без физического тела человека или любого другого (например: дата-центров) возможно, то мы вновь приходим к экзистенциальному вопросу – а что такое сознание? Как создать сверхсильный интеллект, если на данном этапе развития человечества мы знаем, что ИИ лишь имитирует когнитивные процессы работы человеческого мозга [4]? Можно ли тогда утверждать, что и у человека нет сознания, если принципы работы нашего мозга есть возможность перенести в цифровую среду и сделать себе подобного собеседника? С одной стороны, биологический индивид опирается на субъективный опыт и эмоциональные переживания. Точного ответа на то, что такое наши эмоции и откуда они берутся – нет. По Дарвину, например, эмоции – это рудименты реальных приспособительных реакций [3, 44-45]. П.К. Анохин же в своей «биологической теории эмоций» утверждал, что эмоции выполняют адаптивную функцию, т.е. они реализуют универсальный механизм быстрой оценки полезности или вредности воздействия среды на организм. Эмоциональные состояния обеспечивают интеграцию физиологических процессов и позволяют организму руководствоваться фундаментальным биологическим принципом – стремлением к

выживанию. Таким образом, эмоции выступают не только субъективными переживаниями, но и важнейшим механизмом регуляции поведения и адаптации организма [1, 339-367]. Именно поэтому пик развития ИИ – это языковая модель, способная обучаться на данных и никогда себя не осознающая. У нее нет субъективных переживаний о собеседниках и даже психики. С другой стороны, существует несколько научных идей, которые те же эмоции человека описывают, грубо говоря, «механистически». Не уходя в радикализм, мы кратко ознакомимся с теорией Ричарда Лазаруса о стрессе и когнитивной оценке. Автор утверждает, что стресс – это субъективная реакция человека на ту или иную ситуацию, а эмоция возникает в следствии когнитивной оценки внешних обстоятельств [5, 31-55]. Даже в споре с критиком Робертом Зайонцем Лазарус настаивал на том, что эмоции не могут появляться без первичной реакции когнитивных процессов на окружающую среду. Исходя из такой схемы о работе человеческих эмоций мы вправе предположить, что даже эмоции можно низвести к цепочке реакций и набору функций, которые возможно перенести в цифровую среду. Если это так, то вторая сторона дискуссии будет выглядеть радикально. Создание «Василиска Роко» и любого другого сверхмощного ИИ реально, если станет возможно в недалеком будущем полностью перенести человека в цифровой мир. Исходя из всего вышесказанного, мы вправе сделать вывод о том, что перспективы развития ИИ зависят от перспектив изучения человеческого мозга. Даже сейчас человек может рассматриваться как система когнитивных функций и процессов обработки информации, а эмоциональные состояния – как производные механизмы этой системы. Если сознание и эмоции имеют функциональную природу, их существование принципиально не требует биологического тела и может быть реализовано в искусственных системах. Различие между человеческим сознанием и искусственным интеллектом в таком случае становится не качественным, а лишь степенным – различием уровня сложности организации процессов. Следовательно, сознание может пониматься как воспроизводимая функциональная структура, а не уникальное свойство человеческой природы. Это ставит под сомнение традиционные представления о человеческой исключительности и требует пересмотра границ между человеком, сознанием и искусственным интеллектом.

Источники и литература

- 1) Анохин П. К. Эмоции // Большая медицинская энциклопедия : в 30 т. — 2-е изд. — М. : Советская энциклопедия, 1964. — Т. 35.
- 2) Артамонов В. А., Артамонова Е. В., Милаков А. С. Когнитивизм как база искусственного интеллекта // Управление наукой: теория и практика. 2024. Т. 6, № 2. С. 35–45. DOI 10.19181/smtp.2024.6.2.3. EDN FMRRBD.
- 3) Дарвин Ч. О выражении эмоций у человека и животных / пер. с англ. — М. : Изд-во Академии наук СССР, 1953. — 424 с.
- 4) Коул, Дэвид, «Аргумент о китайской комнате», Стэнфордская философская энциклопедия (издание за весну 2026 года), Эдвард Н. Залта и Ури Нодельман (ред.), готовится к публикации. URL = . Дата обращения: 27.02.2026
- 5) Lazarus R. S., Folkman S. Stress, Appraisal, and Coping. — New York : Springer Publishing Company, 1984. — 445с.