

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
УЧРЕЖДЕНИЕ НАУКИ
ИНСТИТУТ ФИЛОСОФИИ И ПРАВА
СИБИРСКОГО ОТДЕЛЕНИЯ РОССИЙСКОЙ АКАДЕМИИ НАУК

СОЗНАНИЕ В ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ

Материалы Международной научной конференции

Новосибирск
2025

ПЛЕНАРНОЕ ЗАСЕДАНИЕ

АНАЛОГОВЫЕ ВЫЧИСЛЕНИЯ И КОМПЬЮТЕРНАЯ МЕТАФОРА СОЗНАНИЯ: ВОЗВРАЩЕНИЕ К ИСТОКАМ

П. Н. Барышников

Пятигорский государственный университет, Пятигорск
pnbaryhnikov@pgu.ru

*Исследование выполнено за счет гранта Российского научного
фонда*

№ 24-28-00540, <https://rscf.ru/project/24-28-00540/>

В этих коротких тезисах не удастся дать полноценную экспозицию проблемы. Это будет своего рода набросок, который ещё предстоит развить в дальнейших исследованиях.

Вопрос, который лежит в основании проблемы звучит следующим образом: почему семантика компьютерной метафоры в когнитивных науках как правило содержит в себе аналогии с электронными цифровыми вычислительными машинами?

В вопросе уже содержится утверждение:

«Семантика компьютерной метафоры в когнитивных науках как правило содержит в себе аналогии с только с электронными цифровыми вычислительными машинами».

История становления Вычислительной теории сознания (Computational Theory of Mind, далее – СТМ) свидетельствует о глубокой методологической (а, точнее, метафорической) взаимосвязи технологического прогресса в области информационных технологий второй половины XX в. и научно-философскими подходами к проблеме сознания. Гипотеза, представляющая сознание и когнитивные процессы как результат вычислений, постепенно стала оформляться в научное направление, взяв в качестве основания утверждение А. Тьюринга о том, что при ограничения перцептивного и когнитивного аппарата любой символический алгоритм, выполняемый человеком, может быть воспроизведен подходящей машиной Тьюринга. Эта идея выражается в следующей формулировке:

«Мы можем сравнить человека в процессе вычисления действительного числа с машиной, которая способна только на конечное число условий $q_1, q_2 \dots q_n$, которые будут называться "т-конфигурациями". Машина снабжена «лентой» (аналогом бумаги), проходящей через нее и разделенной на секции (называемые "квадратами"), каждая из которых способна нести "символ"» [Turing, 1937, p. 231].

Можно утверждать, что с этой формулировки начинается история компьютерной цифровой метафоры, этапы формирования и эвристические функции которой рассмотрены достаточно подробно в обширном корпусе литературы [Baryshnikov, 2023; Richards, Lillicrap, 2022; Boden, 2008]. Причина такого революционного эффекта от абстрактной модели очевидна: Тьюринг изобрёл универсальную машину, которая принципами своей работы способна описать работу любой

другой машины. Возможность имитации любой машины с дискретными состояниями – причина неоспоримой популярности цифрового подхода.

При этом важно отметить, что уже в своих ранних работах Тьюринг разрешил парадокс диагональных процессов, показав, что вычислимые числа могут быть перечислимы, но в действительности их невозможно перечислить за конечное число шагов. То есть этим Тьюринг сразу обозначил ограничение дискретного подхода, доказав, что диагональный аргумент противоречит возможности существования универсального решателя [Boolos et al., 2007]. То есть эти формальные ограничения нельзя было не учитывать. Тем не менее за последние 80 лет общее для большинства эмпирических наук компьютерное моделирование трансформировалось в когнитивных науках в представление о том, что мозг и сознание являются реальными вычислительными системами. Вслед за этой трансформацией за счет неявного метафорического переноса мозг стал наделяться чертами электронной цифровой вычислительной машины. А правила организации ментальных процессов – концептуальными свойствами алгоритмов.

Аналоговый подход предлагает альтернативное видение процесса вычислений. И это видение, если его спроецировать на компьютерную метафору в когнитивных науках, придает новые «семантические смещения», которые обладают, на мой взгляд, не меньшим эвристическим потенциалом, чем цифровой подход.

Эвристический потенциал аналоговой версии компьютерной метафоры основывается на принципиальных отличиях между цифровым и аналоговым видами вычислений (См. Таблицу 1).

Цифровой подход	Аналоговый подход
Информация кодируется через биты (0 и 1), которые формируют числа по позиционному принципу.	Информация не кодируется: величины (например, напряжение, ток, углы поворота, давление) как непрерывные переменные прямо представляют числа : 107 В означает число 107.
Вычислительный, алгоритмический и имплементационный уровни реализованы в разных онтологиях.	Физическая реализация представляет собой одновременно и алгоритм, и структуру.
Программный и аппаратный уровни разведены. Математическая абстракция не совпадает с принципами физической реализации.	Вычисление происходит благодаря механистическим взаимодействиям физических величин.

Таблица 1. Некоторые отличия цифрового и аналогового типов вычислений [Maley, 2024]

Цифровые и аналоговые вычислительные машины обладают базовым общим свойством. Это свойство отличает их от множества других машин. Дело в том, что вычислительные процессы лежат в основании очень многих механизмов (например, автоматическая трансмиссия). Но результат вычислений как результат реализации математической модели в материальной системе является целевой функцией

исключительно вычислительных машин (в том числе и простейших, например, логарифмической линейки). Любая вычислительная машина является аналогией, подобием. То есть машинное вычисление – это перевод математической абстракции на язык физики вычислительных процессов.

Здесь важно отметить следующее: если вычислительная машина преобразует измеримые величины по формальным правилам, и структура вычислительной машины описывается теми же уравнениями, что и материальная система, то это прямая аналогия. Аналоговые вычислительные машины – в большинстве случаев представляют собой прямую аналогию. Элементы одной системы (например, масса, пружина, демпфер в механике) соответствуют элементам другой (например, индуктивность, ёмкость, сопротивление в электричестве). Цифровая же вычислительная машина – это всегда непрямая аналогия. ЦВМ оперируют данными в виде дискретных чисел (0 и 1), а не непрерывных сигналов. Уравнения реальной системы решаются через алгоритмы, а не через физическое соответствие элементов. Например, решение дифференциального уравнения на цифровом калькуляторе требует дискретизации и численных методов, а не прямого отображения физических компонентов.

Большинство процессов в физическом мире аналоговые. Конечная точность измерений – это всегда аппроксимация. Вопрос о цифровой компьютерной метафоре в когнитивных науках звучит все острее. Мозг – сложная физическая материальная система. Аналоговые вычисления оказались непригодными для моделирования когнитивных процессов сразу по нескольким причинам:

- ограниченная точность: возникали ошибки из-за шума, температурных дрейфов и неточностей компонентов;
- непрограммируемость: каждая задача требовала перенастройки схемы;
- цифровая революция: цифровые процессоры предложили универсальность, точность и простоту программирования.

Рассмотрим прямой эвристический эффект аналоговой версии компьютерной метафоры на примере исследования Н. Гарга [Garg, 2025]. В этой диссертации особое внимание уделено проектированию мемристоров. Мемристор – пассивный электронный элемент с переменной проводимостью, которая зависит от свойств протекающего через него заряда. Мемристор можно представить как «резистор с памятью». В рамках нейроморфных систем мемристоры используются как синапсы: их проводимость моделирует силу синаптической связи, а изменение проводимости соответствует обучению (по правилу Хебба: одновременная активация клеток приводит к выраженному увеличению синаптической силы между этими клетками). Н. Гарг утверждает, что аналоговые нейроморфные системы с мемристивными синапсами способны превзойти традиционные цифровые архитектуры по энергоэффективности и производительности в задачах, требующих непрерывного онлайн-обучения (например, обработка данных на периферии, IoT-устройства). Автор обосновывает важность разработки совместимого дизайна алгоритмов и аппаратной части, где физические свойства устройств становятся частью вычислительного процесса.

Говоря о «возвращении к истокам» в заголовке этих тезисов, я имел в виду наличие «аналоговых» рудиментов в

различных современных направлениях цифрового компьютерного моделирования: аналоговые нейросети [Siegelmann, 1999], нейроморфное моделирование и квантовый компьютеринг.

Каждое из упомянутых направлений требует детального анализ со стороны философии когнитивных наук. Если мы исходим из того, что сознание / мозг вычисляет, то необходимо будет вводить корректировки в содержание понятия «вычисление». Если вычисление всегда включает в себя представление и механизм для манипуляции этим представлением, то необходимо найти в когнитивной физической системе корреляты представления. И, конечно, тогда не остается никакого места для цифрового кода и алгоритма оперирования. «Частоту срабатывания нейрона, возможно, лучше рассматривать как представление величины стимула (т. е. аналогового представления, будь то дискретного или непрерывного), а не как поток битов» [Maley, 2024].

Стоит указать на то, что альтернативные вычислительные архитектуры в первую очередь направлены на миниатюризацию устройств и оптимизацию энергетического и вычислительного бюджета, а не на универсализм, который свойственен цифровому подходу. Способно ли аналоговое конструирование привнести прорывные идеи в прикладные когнитивные исследования – вопрос, на который ещё предстоит найти ответ.

Литература

Baryshnikov, P. (2023). Language, Mind, and Computation in the Metaphors of Cognitive Science. *Technology and Language*. 4 (13). P. 1–6. DOI: 10.48417/TECHNOLANG.2023.04.01

Boden, M. A. (2008). *Mind as machine. A history of cognitive science*. Oxford. Clarendon Press.

Boolos, G. S., Burgess, J. P., & Jeffrey, R. C. (2007). Diagonalization. In Boolos, G. S., Burgess, J. P., & Jeffrey, R. C. (ed.) *Computability and Logic*. Cambridge University Press. P. 16–22.

Garg, N. (2025). *Neuromorphic in-memory learning with analog integrated circuits and nanoscale memristive devices* [PhD Thesis, Université de Lille; Université de Sherbrooke]. URL: https://hal.science/tel-04821563v1/file/garg_nikhil_PhD_2024.pdf

Maley, C. J. (2024). Computation for cognitive science: Analog versus digital. *WIREs Cognitive Science*. 15(4). e1679. DOI: 10.1002/wcs.1679

Richards, B. A., Lillicrap, T. P. (2022). The Brain-Computer Metaphor Debate Is Useless: A Matter of Semantics. *Frontiers in Computer Science*. 4. 810358. DOI: 10.3389/fcomp.2022.810358

Sieglmann, H. T. (1999). *Neural Networks and Analog Computation*. Boston. Birkhäuser. DOI: 10.1007/978-1-4612-0707-8

Turing, A. M. (1937). On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*. 42(1), 230–265. DOI: 10.1112/plms/s2-42.1.230

АГЕНТНЫЙ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И ВИРТУАЛЬНАЯ РЕАЛЬНОСТЬ

Д. И. Свириденко

Институт философии и права СО РАН, Новосибирск
dsviridenko47@gmail.com

В данной работе обсуждаются текущая и потенциальная роль и значение агентного искусственного интеллекта (ИИ) в становлении и развитии виртуальной реальности (VR, искусственная действительность), понимаемой как технология, позволяющая пользователям с помощью специальных устройств, способных исключительно полно и правдоподобно имитировать взаимодействие с виртуальной средой путём воздействия на все пять имеющихся у человека органов чувств, погружать пользователя в созданный цифровой мир, менять его подход к пониманию и решению различных задач и взаимодействию с окружающим миром. При этом для создания иммерсивности среды генерация свойств и реакций виртуальной реальности производится в максимально возможном согласии с реальными законами физики (гравитация, свойства воды и воздуха, столкновение с предметами, отражение и т.п.) и в реальном времени. Заметим, что пользователю виртуального мира, часто в маркетинговых или развлекательных целях, предоставляется больше возможностей, чем в реальной жизни (летать, стрелять, бегать с высокой скоростью, высоко прыгать, нырять и плавать под водой, создавать предметы и т.п.).

Следует отметить, что современный ИИ уже значительно преобразил виртуальную реальность, упростив и ускорив процесс создания контента (декларативная и процедурная генерация контента, а также его оптимизация, автоматизация проектирования и кодирования), предоставляя новые возможности виртуального моделирования (адаптивные симуляции, реалистичные сценарии, предиктивная аналитика и т.п.), что позволило к настоящему времени существенно улучшить пользовательский опыт (персонализация, естественность взаимодействия, адаптация в реальном времени).

Что же касается будущего концепции виртуальной реальности, то, по мнению автора, наибольшее влияние на ее дальнейшее развитие будет оказывать, прежде всего, агентный ИИ, выступая в роли основного инструментария. И начало тому уже положено – в виртуальной реальности начали активно внедряться и использоваться так называемые ИИ-агенты, что позволяет адекватно моделировать поведение и осуществлять персонализацию действующих в VR акторов. К настоящему времени постепенно складывается и типовая схема работы не только отдельного ИИ-агента, но и их сообществ: пользователь ставит перед главным ИИ-агентом задачу, ИИ-агент планирует, распределяет и выполняет работу, разбивая рабочий процесс на подзадачи, которые агент-менеджер назначает другим специализированным субагентам, координирующим работу друг с другом (при этом сообщество ИИ-агентов может итеративно улучшать вывод, запрашивая у пользователя дополнительный ввод данных для обеспечения точности и релевантности), система ИИ-агентов выполняет действия, необходимые для

полного завершения решения задачи. Развивается и классификация ИИ-агентов:

- индивидуальное дополнение/ИИ-ассистент/ИИ-помощник/агент «второго пилота»/ко-пилот;
- ИИ-платформы автоматизации рабочих процессов;
- (Gen AI-native)-агенты для доменных решений – системы обслуживания клиентов, конвейеры разработки программного обеспечения и т.п.;
- (AI-native)-агенты предприятия и операционные модели;
- виртуальные работники ИИ – ИИ-агенты, которые функционируют как сотрудники или члены команды и представляют собой наиболее перспективную и бурно развивающуюся категорию ИИ-агентов.

В настоящее время сам агентный ИИ становится центральным направлением развития ИИ, все больше приобретая черты гибридного ИИ – подхода, целью которого является создание таких моделей вычислимости, которые позволяли бы эффективно сочетать достоинства и преимущества как символического ИИ, так и машинного обучения. В докладе предлагается различать технологический вариант интеграции направлений ИИ, базирующийся на процедуре векторизации, что позволяет, например, граф знаний вкладывать в нейронную сеть, и методологический вариант интеграции, примером которого может служить агентный подход, развиваемый в [Стюарт Рассел, Питер Норвиг, 2006]. Здесь следует отметить не только предлагаемую агентным подходом развитую методологию и технологию интеграции направлений ИИ, но и оригинальность трактования самого понятия «искусственный интеллект» как науки о рациональных агентах, т.е. как науки о некоторых

устройствах или программах, способных воспринимать среду своего существования и выполнять действия, максимизирующие их шансы на успех при достижении какой-то цели.

Другим примером интеграции символьного ИИ и машинного обучения, развивающий и обобщающий концепцию агентного подхода, является задачный подход к ИИ [Goncharov et al., 2023] и его практическая реализация в виде нескольких инструментально-технологических платформ и языков исполнимых спецификаций задач, позволяющих по объяснительной и интерпретируемой спецификации задач генерировать исполнимый код, приводящий к их решениям. Спецификации задач, относящихся к некоторой предметной области, по которым формируется соответствующий запрос к предметной области, поиск ответа на запрос и получение решения (ответа на запрос) осуществляется в рамках и средствами представленного в работе концепции семантического моделирования [Гончаров, Свириденко, 1986; Goncharov, Sviridenko, 1987; Goncharov, Sviridenko, 1989; Гончаров, Свириденко, 2018], являющейся, фактически, логико-математической теорией задачного подхода. Основу этой теории составляет так называемая семантическая вычислимость, сочетающая в себе формульную определимость [Ершов Ю.Л., 2000], оракульную (относительную) вычислимость и обучаемые семантические модели. Отметим, что именно оракульная вычислимость и позволяет создавать не только системы гибридного ИИ общего вида [Goncharov S., Sviridenko D., 2019], но и формальные модели специализированных рациональных агентов гибридного типа и их сообществ – мультиагентных систем, имеющих динамическую природу. С этой целью в

докладе рассматривается расширенный вариант семантического моделирования, носящий название семантического динамического моделирования, языковыми средствами которого и предлагается осуществлять формальную спецификацию динамического поведения гибридных рациональных агентов. В докладе также обсуждаются возможные полезные обогащения построенного формализма динамической логики дополнительными неклассическими логическими средствами – модальностями, темпоральными операторами, кванторами путей и т.п. В заключение ставится задача обучения гибридных рациональных агентов и обсуждаются варианты ее решения средствами вероятностного семантического вывода [Витяев Е.Е., 2017].

Литература

Витяев, Е. Е. (2017). Семантический вероятностный вывод предсказаний. *Известия Иркутского государственного университета, Серия «Математика»*. Том 21. С. 33–50.

Витяев, Е. Е., Гончаров, С. С., Свириденко, Д. И. (2020). О задачном подходе в искусственном интеллекте и когнитивных науках. *Сибирский философский журнал*. Т. 18, № 2. С. 5–29.

Гончаров, С. С., Свириденко, Д. И. (1986). Математические основы семантического программирования. *Доклады АН СССР*. 289: 6. Р. 1324–1328.

Гончаров, С. С., Свириденко Д. И. (2018). Семантическое моделирование и искусственный интеллект. *Сибирский философский журнал*. Т. 16, № 4. С. 5–25.

Ершов, Ю. Л. (2000). *Определимость и вычислимость*. Новосибирск.

Рассел, С., Норвиг, П. (2006). *Искусственный интеллект: современный подход*. Киев, Вильямс, 1409 с.

Ershov, Yu. L., Goncharov, S. S., Sviridenko, D. I. (1987). Semantic foundations of programming. *Lecture Notes in Computer Science*. Vol. 278. P. 116–122.

Goncharov, S. S., Sviridenko, D. I. (1989). Σ -programming. Transl., II. Ser., *Am. Math. Soc.*, Vol. 142. P. 101–121.

Goncharov, S., Sviridenko, D. (2019). Semantic Modeling and Hybrid Models. *2019 International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON)*, Novosibirsk, Russia. P. 987–990.

Goncharov, S. S., Vityaev, E. E. Sviridenko, D. I. (2020). Problem-solving approach in artificial intelligence. *Applied Mathematics and Fundamental Informatics*. T. 7. No. 2. P. 4–9.

Vityaev, E.E., Goncharov, S.S., Sviridenko, D.I. (2023). Task-driven approach to artificial intelligence. *Cognitive Systems Research*. Volume 81, September 2023. P. 50–56.

Быть или казаться?

А. Ю. Сторожук, А. С. Зайкова

ИФПР СО РАН, Новосибирск

stor71@mail.ru, zaykova.a.s@gmail.com

Тест Тьюринга, или имитационная игра, позволяет говорить о возможности имитации человеческого поведения машиной. Этот тест был предложен для обхождения вопроса о сути мышления (think). Алан Тьюринг предполагал, что, в случае, когда квалифицированный наблюдатель не сможет отличить ответы машины от машины человека, разумнее всего считать машину мыслящей.

Однако за этой прагматической формулировкой скрывается глубокий метафизический вопрос о метафизическом тождестве оригинала и имитации: может ли тождество поведения служить основанием для отождествления сущности? Если в некоторых определенных условиях объект А ведет себя сходно с объектом Б, является ли тождество поведения основой для метафизического отождествления?

Философское обоснование теста Тьюринга может выступать как одно из следствий подхода структурного, или онтического реализма [Beni, 2016]. Данная позиция принимает ряд тезисов, которые отвергают метафизическую значимость сущностей. В частности, в парадигме онтического реализма можно сформулировать следующие утверждения:

1. Элиминативизм: индивидов не существует (но существует структура отношений).

2. Существуют отношения (или реляционные факты), которые не зависят от внутренних и пространственно-временных свойств их объектов.

3. Отдельные объекты не имеют внутренней природы [Ainsworth, 2011].

4. Существуют отдельные объекты или вещи, но они не обладают какими-либо неотъемлемыми внутренними свойствами [French, Ladyman, 1999].

Таким образом, если некие различные объекты совпадают в своих проявлениях, с точки зрения онтического реализма их сущности тождественны. Эта концепция распространилась достаточно широко, поскольку под объектами имелись ввиду сущности микромира, для которых справедлив принцип тождества элементарных частиц. К примеру, два электрона в одинаковых состояниях полностью идентичны.

Но когда речь идет о сознании и индивидуальности, онтический реализм вряд ли найдет столь же много сторонников, как в случае с фермионами. Наличие самосознания позволяет каждому субъекту отличать себя от всего остального мира и верить в свою уникальность. Даже если два субъекта будут вести себя одинаково, мы не можем быть уверены, что их внутренний опыт совпадает.

В качестве наглядного примера можно вспомнить эксперимент по воспитанию детёныша шимпанзе вместе с ребёнком, проведённый в 1930-х Уинтропом и Луэллой Келлогг. Семья зоологов выращивала совместно со своим ребенком детёныша шимпанзе Гуа [Kellogg, 1931]. Шимпанзе Гуа научилась одеваться, есть ложкой, выучила язык жестов, то есть

имитировала человеческое поведение. Но стала ли обезьяна человеком? Если бы это было так, воспитанием можно было бы менять сущность индивида и мы близки к пути Лысенко.

Аналогично, если машина успешно проходит тест Тьюринга, это не доказывает, что она мыслящая в том же смысле, что и человек. Поведенческое тождество в условиях теста не гарантирует внутреннего тождества во всех возможных обстоятельствах. Поэтому нам следует принять сущностное онтологическое отличие, (отрицать вышеперечисленные пп. 1–4) и в вопросе сравнения человека с машиной, даже если имитация окажется очень похожей. Иными словами, мы должны признать наличие скрытых существенных качеств индивидов и их существенность.

Остается вопрос, почему мы считаем тождество метафизической природы более важным, чем тождество, основанное на совпадении внешних проявлений. Ответ на этот вопрос лежит в большей устойчивости метафизических сущностей, чем операционального тождества. Представим себе человека и машину, которая его имитирует. В обычной среде они могут действовать одинаково. Но что будет, если отключить электроэнергию? В изменившихся условиях человек будет и дальше действовать как человек, в то время как поведение машины сильно изменится. Поскольку невозможно предусмотреть все возможные ситуации, следует признать ограниченное значение теста Тьюринга.

Таким образом, тест Тьюринга полезен как инструмент выявления определённого уровня поведенческого интеллекта, но он недостаточен для ответа на вопрос о возможности отождествления сущности. Остаётся открытым вопрос о том,

может ли существовать способ проверки не только внешних, но и внутренних характеристик системы. Даже если бы такой тест мог бы включать проверку способности субъекта адаптироваться к радикально изменённым условиям, сохранять идентичность в долгосрочной перспективе, обладать автономной целевой активностью, независимой от внешнего программирования, это не гарантировало бы возможность или допустимость онтологического тождества человека и машины.

Литература

Ainsworth, P. (2011). Ontic Structural Realism and the Principle of the Identity of Indiscernibles. *Erkenntnis*. N. 75. P. 67–84. DOI: 10.1007/s10670-011-9279-x

Beni, M. D. (2016). Epistemic Informational Structural Realism. *Minds and Machines*. № 26 (4). P. 323–339. DOI: 10.1007/s11023-016-9403-4

French, S. Ladyman, J. (1999). “Reinflating the Semantic Approach”, *International Studies in the Philosophy of Science*, № 13(2). P. 103–121. doi:10.1080/02698599908573612

Kellogg W. N. (1931). HUMANIZING THE APE. *Psychological Review*. № 38. P. 160–176.

Машина мысли:
эволюция компьютерной метафоры сознания
в эпоху больших вычислений

ПИФАГОРЕЙСКОЕ УЧЕНИЕ И АРХИТЕКТОНИКА СИМУЛЯЦИИ: НЕОЧЕВИДНЫЕ СВЯЗИ

А. Н. Габов

Институт философии и права СО РАН, Новосибирск
realgabov@gmail.com

Известно, что учение пифагорейцев крайне тяжело поддается строгому и окончательному формулированию по причине достаточно сильно выраженной разнородности составляющих его векторов. В зависимости от модальности мысли той или иной группы внутри школы Пифагора вопрошания о природе реальности, как правило, получали различные, порой диаметрально противоположные версии ответов. Не последнее место в списке таких ответов занимает так называемое «тайное» или «сокровенное» учение пифагорейской школы.

Основной тезис настоящей статьи заключается в том, что ключевые компоненты сокровенного учения пифагорейцев, изложенные в диалоге Платона «Федон» во многом схожи с догадками Ника Бострома относительно возможной архитектоники симуляции, в которой мы гипотетически обитаем.

1. Первичность и удвоение

Мысль о том, что видимому нами консенсусу реальности *предлежит* некая первичная реальность, критически значимая для существования нашего консенсуса, но в той или иной мере сокрытая от обыденной оптики, является одной из наиболее важных проблем раннегреческой философии. Мысль эта, несомненно, корнями уходит в мифологические представления греческой цивилизации. Например, *предлежание* появления Олимпийских богов появлению упорядоченного космоса есть помимо прочего критический пункт логики о вторичности мира как таковой. Этот процесс можно условно назвать *нисходящим* удвоением (принцип: как наверху – так и внизу, яко на небеси – тако и на земли). Вместе с тем, попытки человека открыть и постичь эту первичную реальность, как правило, приводят к *восходящему* удвоению. В этих рекурсивных отражениях порой возникают причудливые картины. О некоторых из них мы и поговорим в этой статье.

2. Диалог «Федон». Пассаж о стражах

В платоновском диалоге «Федон» есть известный фрагмент, где упоминается пифагорейское учение. Речь в данном фрагменте идет о непозволительности самоубийства. В частности, на вопрос Кебета о том «почему же все-таки... считается, что убить самого себя непозволительно?» Сократ отвечает, что «среди всего прочего лишь это одно так просто и не терпит никаких исключений, как бывает во всех остальных случаях» [Платон, 1993, с. 12]. При этом Сократ даёт следующую аргументацию: «Сокровенное учение гласит, что мы, люди, находимся как бы под стражей и не следует ни избавляться от неё своими силами, ни бежать, – величественное, на мой взгляд, учение и очень

глубокое... о нас пекутся и заботятся боги, и потому мы, люди, – часть божественного достояния» «Сокровенное учение», согласно комментариям Лосева, это учение пифагорейцев [там же, с. 430]. Сократ продолжает: «Но если бы кто-нибудь из тебе принадлежащих убил себя, не справившись предварительно, угодна ли тебе его смерть, ты бы, верно, разгневался и наказал бы его, будь это в твоей власти?» [Платон, 1993, с. 12]. Как усматривается из приведенного отрывка, Платон, по всей видимости, воспроизводит самую суть пифагорейского учения. Дадим представление относительно основных компонентов этого учения.

3. Краткая формулировка пифагорейского «сокровенного учения» на основе диалога «Федон»

Метафизика «сокровенного учения» состоит из следующих неотъемлемых компонентов:

1. Нахождение всей совокупности людей под неким контролем, причем: а) никому не следует произвольно избавляться от контроля; б) не следует бежать; в) под «бегством» явно подразумевается, исходя из контекста данного фрагмента, не что иное, как самоубийство; г) контроль со стороны богов подразумевает «опеку» и «заботу».

2. Учение, по мнению Сократа, является величественным и глубоким, то есть проникает в самую глубину структур нашей реальности.

3. Мы принадлежим неким неназванным богам, мы являемся их «достоянием», фактически – собственностью. Таким образом, запрет на самоубийство, по всей видимости, связан с произвольной порчей «собственности» богов.

4. Возможная архитектура симуляции по Нику Бострому

Бостром в статье “*Are you living in a computer simulation?*” (*Живём ли мы в компьютерной симуляции?*) даёт описание характеристик симуляции, в которой мы можем находиться:

1. Симуляция может быть порождена некоей базовой постчеловеческой цивилизацией, чьи компьютеры достигли необходимого уровня вычислительных мощностей. Важным уточнением является при этом тот факт, что «физика вселенной, где находится компьютер, может напоминать, а может и не напоминать физику наблюдаемого нами мира» [Bostrom, 2003, p. 253].

2. Рекурсивность симуляции. Базовая по отношению к нашей симуляции цивилизация может быть не конечной, но также лишь симуляцией вышестоящей цивилизации более высокого уровня [там же].

3. Постлюди базовой по отношению к нашей симуляции цивилизации «подобны богам». Характерно, что по Бострому «...они всемогущи в том смысле, что они могут вмешиваться в работу нашего мира способами, нарушающими физические законы...» [там же, p. 254].

4. Ограничения для обитателей симуляции связано с этическим компонентом. Автор полагает, что действия на уровне обитателей симуляции «могут повлиять на отношение к ним обитателей более глубокого уровня реальности. Например, если никто не может быть уверен, что он находится на базовом уровне, то каждый должен рассматривать вероятность того, что его действия будут вознаграждаться или наказываться, возможно, на основании неких моральных критериев, хозяевами симуляции» [там же].

5. Ограничение для обитателей симуляции связано с дефицитом понимания мотивации, целеполагания и мышления постлюдей базовой цивилизации [там же, р. 255].

5. Неочевидные связи

Стражи или боги, чьим достоянием, по Платону, мы, люди, являемся, функционально («пекутся и заботятся») напоминают представителей базовой по отношению к нашей гипотетической симуляции цивилизации постлюдей. Нахождение обитателя симуляции под контролем и в том, и в другом случае. Технически базовая цивилизация способна «выключить» симуляцию целиком или отключить обитателя симуляции, равно как и боги в платоновском варианте способны лишить человека жизни или в целом уничтожить наш мир.

Проблема попыток обитателя симуляции избавиться от контроля, сбежать из симуляции, пренебрегая «опекой» базовой цивилизации, прямо не описана Бостромом. Однако, исходя из очевидно конструктивного целеполагания постлюдей в отношении симуляции, можно предположить, что такое поведение с большой долей вероятности не будет вписываться в нормальный, предполагаемый ими, ход течения вещей. Частный случай платоновской проблемы запрета на самоубийство в данном случае может быть раскрыт в виде вопроса о нарушении обитателем симуляции права собственности постлюдей на всё содержимое симуляции, включая самих её обитателей. Кроме того, такое недопустимое поведение может нарушать некие сценарии, заранее прописанные для того или иного блока симуляции, влиять на конечные результаты, на стабильность программы в компьютере базовой цивилизации; уже в силу этого

оно должно осуждаться богами-постлюдьми в той же мере, в которой самоубийство осуждается в сокровенном учении пифагорейцев.

Отдельного рассмотрения заслуживает вопрос пифагорейского запрета на распространение «сокровенного» учения и возможный «программный» запрет базовой цивилизации на обсуждение параметров и создателей симуляции её обитателями.

Литература

Платон. (1993). Собрание сочинений в 2 т. Т. 2. М.: Мысль. 528 с.

Bostrom, N. (2003). Are you living in a computer simulation? *The Philosophical Quarterly*. Vol. 53. № 211. P. 243–255.

Применимость эпистемических теорий в решении проблемы галлюцинирования LLM

В. О. Дугушкин, В. И. Кочергин

Новосибирский государственный университет

Новосибирск

Крупные языковые модели (LLM), такие как DeepSeek и ChatGPT, в силу их стремительного развития как инструмента с широким набором востребованных функций для глобального пользователя, стали символом прогресса в области обработки и генерации естественного языка. Популярность и широкая применимость данных моделей выходят далеко за пределы академических исследований, однако именно это стремительное распространение технологий вызывает настороженность и даже определённое недоверие у части философской публики (Daniel Dennett, Timnit Gebru, Emily Bender). Нередко, среди причин «недоверия» к LLM одной из ключевых является, так называемое, «галлюцинирование» – выдача не соответствующих реальности, логически противоречащих или полностью сфабрикованных по содержанию утверждений [Cleti, Jano, 2024, p. 1].

К настоящему моменту, решение проблемы галлюцинирования является одним из основных направлений исследований в области ИИ. Попытки решения этой проблемы уже предпринимались, в том числе с использованием эпистемологических теорий (например, [Sekrst, 2024]), однако такие попытки сводились в первую очередь к оценке убеждений, сформированных у пользователя посредством взаимодействия с LLM. Мы же полагаем, что приписывание эпистемических

статусов может производиться и самой LLM: эпистемологические термины могут быть применены для создания необходимых ориентиров (framework), соответствие которым будет предположительно повышать надежность модели. Заданные ориентиры, в свою очередь, могут быть транслированы модели с помощью специально сконструированных промтов (запросов, которые пользователи передают нейросети для решения нужной пользователю задачи) или встроены в сам процесс обучения.

Таким образом, цель нашей работы заключается в том, чтобы предложить философские основания решения проблемы галлюцинирования LLM. Мы исходим из того, что специфика работы подобных систем предполагает ограниченный доступ к их внутреннему устройству, а значит, и их оценка может осуществляться только на основе внешних данных — входных (input) и выходных (output). При этом сама «непрозрачность» устройства модели открывает возможность для её медиации через введение инструментальных понятий и концептуальных ориентиров, которые придадут работе модели дополнительный уровень интерпретационной структуры.

Характерной особенностью LLM является восприятие всей базы обучающей информации как равным образом истинной [He, Yang, 2025, p. 3–4]. К примеру, информация «цвет луны – серый», полученная из научно-ориентированного источника, и информация «цвет луны – красный», описывающая личный опыт наблюдения и полученная из иного источника, считаются как равноценные, при условии, что оба блока данных были использованы при формировании базовой модели.

Мы полагаем, что все положения, принятые моделью в процессе обучения, могут быть проинтерпретированы как базовые убеждения (basic beliefs). Таким образом, можно представить процесс генерации как выведение новых положений из базовых с сохранением истинности (как последняя понимается в свете когерентного подхода к истинности). Такая схема работы LLM соответствует интуициям фаундалистской теории познания.

Тем самым, нами предложено применение фаундализма в качестве философского основания решения проблемы галлюцинирования LLM. Подобное решение предполагает не только формирование цельной системы концептуальных ориентиров, но и разработку способов их интеграции в практику использования языковых моделей. Это, в свою очередь, открывает возможность присваивания эпистемических статусов отдельным «убеждениям» модели с опорой на понятийный аппарат интерналистских теорий эпистемического обоснования, что позволяет рассматривать данную стратегию как перспективное направление в развитии философии искусственного интеллекта.

Литература

Cleti, M., Jano, P. (2024). *Hallucinations in LLMs: Types, Causes, and Approaches for Enhanced Reliability*. DOI: 10.13140/RG.2.2.12184.61445

He, J., Yang, C. (2025). Testimony by LLMs. *AI & Society*. DOI: 10.1007/s00146-025-02366-y

Sekrst, K. (2024). Chinese Chat Room: AI Hallucinations, Epistemology and Cognition. *Studies in Logic, Grammar and Rhetoric*. 2024. №69 (82). P. 370–375.

ПРИЧИННО-СЛЕДСТВЕННЫЕ ОТНОШЕНИЯ В ИССЛЕДОВАНИЯХ ЕСТЕСТВЕННОГО И ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

С. Е. Ильин, А. В. Ильина

Санкт-Петербургский государственный электротехнический
университет «ЛЭТИ» им. В.И. Ульянова (Ленина), Санкт-
Петербург
s.ilin2086@gmail.com

Излагая основания своей философской системы, французский мыслитель Р. Декарт избирает отправным пунктом рассуждения тезис, в истинности которого не может сомневаться: «я мыслю, следовательно, я существую» [Декарт, 2019, с. 52]. Выбранная формулировка строится как логическое следование и актуализирует вопросы о природе связи между существованием и мышлением, а также о механизмах, которыми обеспечивается способность человека мыслить. В век бурного развития искусственного интеллекта (далее – ИИ) эти вопросы обретают новое прочтение при решении задачи о реализации мышления в машинах. Разработка мыслящих машин оказывается сегодня одним из путей более глубокого понимания машинерии человеческой мысли.

Достижения в области ИИ оказывают существенное влияние на жизнь людей и практики функционирования социальных институтов. Однако, в то же время, фиксируется ряд проблем, связанных с применением ИИ. Среди прочего, речь идет об угрозе безопасности обрабатываемых данных, фактах

предвзятости ИИ, а также о принятии решений по принципу «черного ящика», т.е. в отсутствие возможности прояснить их основания. В целях преодоления перечисленных проблем исследователи предлагают концептуальные подходы к совершенствованию ИИ, включая ответственный ИИ (англ. Responsible AI), объяснимый ИИ (англ. Explainable AI) и др. Важное место среди них занимает т.н. каузальный ИИ (англ. Causal AI), в рамках которого от моделей ИИ требуется обнаруживать причинно-следственные связи между явлениями [Cavique, 2024].

При разработке каузального ИИ особый интерес представляют исследования на стыке когнитивной нейробиологии и собственно ИИ. Первые вносят вклад в анализ связей между когнитивными процессами и активностью головного мозга, тогда как вторые – в изучение и разработку методов моделирования этих связей в искусственных нейронных сетях. В результате создаются условия для взаимного обогащения каждой из дисциплин новыми наблюдениями. Обращение к когнитивной нейробиологии позволяет применить знания о естественном интеллекте в условиях работы над ИИ, в то время как работа над ИИ открывает новые возможности для изучения функций естественного интеллекта. Междисциплинарный характер обозначенной проблематики, а также необходимость применения общенаучных методов познания делают соответствующие исследования перспективным предметом философского рассмотрения.

Обсуждение указанных исследований представляется возможным организовать вокруг двух основных вопросов. С одной стороны, вопроса о том, как следует определять понятие

«причинность» (англ. causality) в контексте изучения биологических нейронных сетей. С другой стороны, вопроса о том, какое применение допускают причинно-следственные отношения в условиях работы импульсных нейронных сетей (англ. spiking neural networks).

К настоящему моменту анализ причинно-следственных отношений в рамках нейронаук успел получить распространение. Как результат, на базе обозначенных дисциплин сформировалось несколько подходов к использованию понятия «причинность». Часть из них допускает возведение к философским определениям причинности, в частности, к определениям через категории производства и контрфактической зависимости в интерпретации американского философа Н. Холла (N. Hall) [Barack et al., 2022]. Используя первую категорию, причины можно определить как события в мозге, которые вызывают другие события. Данный подход располагает к избирательному анализу причинно-следственных отношений. При обращении к нему перспективно игнорировать события, включенные в причинно-следственную цепь, но не являющиеся непосредственными причинами изучаемых феноменов. Например, если речь идет о реакции страха в ответ на резкий звук, причиной этой реакции имеет смысл считать работу амигдалы, а не уха, которым звук был воспринят [Barack et al., 2022, p. 654]. Альтернативный подход подразумевает рассмотрение причин как факторов, от которых зависят исследуемые события и феномены. В этом случае факторы, которые оказывают на предмет изучения опосредованное влияние, остаются релевантными, поскольку расширяют диапазон стратегий по контролю его состояния. В примере с реакцией страха в ответ на резкий звук это означает

сохранение внимания к работе уха при восприятии звукового сигнала. Наряду с названными подходами к определению понятия причинности в нейронауках могут быть выделены другие. Как результат, перечень может быть продолжен [Varack et al., 2022, p. 654–655].

Упомянутые подходы к интерпретации причинности объединяет тот факт, что они ориентированы на поиск причинно-следственных отношений в данных о функционировании биологических нейронных сетей. Иными словами, они проецируют представления о причинности на биологические нейронные сети для лучшего их понимания. Обнаружение причинно-следственных отношений выступает целью исследования или средством достижения поставленной цели. В работах, посвященных обсуждению импульсных нейронных сетей, акценты оказываются расставлены несколько иначе.

Импульсные нейронные сети представляют собой один из вариантов искусственных нейронных сетей наряду с архитектурами на базе глубокого обучения. В то время как последние работают с упрощенными моделями нейронов, первые призваны моделировать функционирование биологических нейронов и нейронных сетей с высокой степенью правдоподобия. При их проектировании предлагается принимать во внимание особенности кодирования сигналов в мозге, строение мембраны нервных клеток, роль нейромедиаторов в обработке информации и т.д. [Yamazaki et al., 2022]. Исследователи, вовлеченные в работу над темой, чаще фокусируются на практической реализации алгоритмов, которые способны обеспечить анализ информации в условиях обмена нервными импульсами. Среди прочего, это оказывается справедливо для анализа причинности.

Речь идет, по меньшей мере, о двух аспектах темы. Во-первых, об использовании импульсных нейронных сетей для выявления причинно-следственных отношений в обрабатываемых данных. Исследуются сценарии применения импульсных сетей для решения практических задач; производится оценка скорости работы сетей и качества получаемых результатов [Moreno-Vote, Drugowitsch, 2015]. Во-вторых, категория причинности привлекается при обсуждении алгоритмов обучения импульсных нейронных сетей. Конкретнее, она используется при корректировании выраженности связей между нейронами. Выраженность связей изменяется в соответствии с вкладом нейронов в правильный или ошибочный ответ на стимул, что позволяет закреплять успехи, избегать повторения ошибок и повышать качество работы сети [Lansdell, Kording, 2023].

Обобщая сказанное, можно сделать следующие выводы. Понятие «причинность» получило обсуждение в исследованиях как биологических, так и искусственных нейронных сетей. Оно располагает разными интерпретациями и допускает применение в теоретических и прикладных исследованиях. Работа с причинно-следственными отношениями открывает дополнительные возможности обращения к философским основаниям когнитивных наук и входит в число стимулов для рефлексии над понятийно-категориальным аппаратом исследований в соответствующей области. Кроме того, анализ причинно-следственных отношений выступает важным направлением разработок в области совершенствования ИИ, что закрепляется на уровне концептуального подхода к созданию каузального ИИ.

Перефразируя Р. Декарта, можно сказать, что, если ИИ мыслит, он существует. Анализ причинно-следственных отношений входит в число важнейших проявлений мышления человека. Реализация соответствующего функционала на базе ИИ расширит его мыслительные возможности и укрепит его экзистенциальный статус.

Литература

Декарт, Р. (2019). *Рассуждения о методе; Начала философии; Страсти души*. Пер. с фр. М. Эксмо.

Barack, D. L. et al. (2022). A Call for More Clarity around Causality in Neuroscience. *Trends in Neurosciences*. Vol. 45. No. 9. P. 654–655. DOI: 10.1016/j.tins.2022.06.003.

Cavique, L. (2024). Implications of Causality in Artificial Intelligence. *Frontiers in Artificial Intelligence*. Vol.7. P. 1–5. DOI: 10.3389/frai.2024.1439702.

Lansdell, B. J., Kording K. P. (2023). Neural Spiking for Causal Inference and Learning. *PLoS Computational Biology*. Vol. 19. No. 4. P. 1–25. DOI: 10.1371/journal.pcbi.1011005.

Moreno-Bote, R., Drugowitsch, J. (2015). Causal Inference and Explaining Away in a Spiking Network. *Scientific Reports*. Vol. 5. P. 1–18. DOI: 10.1038/srep17531.

Yamazaki, K., et al. (2022). Spiking Neural Networks and Their Applications: A Review. *Brain Sciences*. Vol. 12. No. 863. P. 1–30. DOI: 10.3390/brainsci12070863.

МОДЕЛИ СОЗНАНИЯ МЕТАСУБЪЕКТОВ ТРЕТЬЕЙ ИСКУССТВЕННОЙ ПРИРОДЫ: РЕКУРСИВНЫЙ ПОДХОД

А. А. Кузнеченков

Самарский университет, Самара

ankuzn1970@yandex.ru

«В XXI веке техническое развитие определяется технологиями искусственного интеллекта (ИИ), задачей создания автономных машин, способных независимо от человека соединять в себе действие машин по переработке энергии ("первая искусственная природа") и по переработке информации ("вторая искусственная природа"), т.е. выполнять без человеческого участия функции познания и технической деятельности. Это переход к третьей искусственной природе, среде обитания, в которой появляются интеллектуальные машины, выполняющие за человека работу по рефлексии» [Нестеров А. Ю., Дорошин А. В., Никоноров А. В., Сойфер В. А., 2022]. Приобретение человечеством состояния третьей искусственной природы напрямую связан с развитием компьютерных систем и рефлексивным аспектом моделей. «Математическая логика и теория алгоритмов, оформившиеся в полной мере в ходе тесного контакта с философскими поисками, легли в основу важнейшего технологического прорыва, а именно, создания современных компьютеров» [Целищев В. В., 2021], что определяет логико-алгоритмическую природу моделей второй и третьей искусственной природы.

Кибернетический подход обеспечивает междисциплинарную связь исследования, позволяет использовать результаты, полученные в различных областях кибернетического знания – от социальной кибернетики до кибернетики технических систем. Развитие кибернетического знания позволяет рассматривать понятие «кибернетика автоматов» для разграничения в рамках технической кибернетики и определения кибернетических областей моделей второй и третьей искусственной природы. Значимым результатом в исследованиях сложных саморазвивающихся кибернетических систем представляют работы В. Е. Лепского, в которых в контексте развития представлений о научной рациональности [Стёпин, 1989] определены этапы становления субъектно-ориентированного подхода [Лепский, 2014]. В области саморазвивающихся систем введено понятие кибернетики третьего порядка [Лепский, 2022] с формированием парадигмы «субъект–метасубъект». Метасубъектный характер кибернетики третьего порядка соответствует теории метасистемных эволюционных переходов В. Ф. Турчина [Турчин, 2000]. Результаты, полученные Лепским, позволяют рассматривать категорию «метасубъект» в качестве основной в моделях третьей искусственной природы, которая напрямую связана с рефлексивными моделями сознания [Лефевр, 1973].

Развитие рекурсивной парадигмы [Хуэй, 2020], исследования рефлексивного аспекта гёделевского аргумента [Целищев, 2021] позволяют использовать рекурсивный подход в исследованиях метасубъектности, в качестве основной характеристики категории метасубъект. Рассматривается предположение о том, что именно рекурсивный дуализм

«субъект-метасубъект» в моделях третьей искусственной природы может стать основой для реализации рефлексивных аспектов моделей сознания, основанных на построении модели окружающего мира и субъекта в нём [Лефевр, 1973]. Субъектно-метасубъектная парадигма наиболее полно отражает эволюционный характер моделей сознания. Рефлексивный аспект гёделевского предложения [Целищев, 2021] позволяет рассматривать понятие «гёделева полнота» в эволюционном контексте, как основу для развития рефлексивных моделей третьей искусственной природы.

Исследования содержания категории «метасубъектность», как основной характеристики метасубъекта, в рамках рекурсивного субъектно-метасубъектного дуализма позволяют определить подходы к построению рефлексивных и саморазвивающихся моделей сознания искусственных метасубъектов в рамках третьей искусственной природы.

Литература

Лепский, В. Е. (2014). Становление субъектно-ориентированного подхода в контексте развития представлений о научной рациональности. *Наука и социальная картина мира. К 80-летию академика В.С. Степина*; под ред. В. И. Аршинова, И.Т. Касавина. М. С. 392–420.

Лепский, В. Е. (2022). Философско-методологические основания постнеклассической кибернетики третьего порядка. *Вопросы философии*. № 8. С. 211–215.

Лефевр, В. А. (1973). *Конфликтующие структуры*. Издание второе, переработанное и дополненное. М.

Нестеров, А. Ю., Дорошин А. В., Никоноров А. В., Сойфер В. А. (2022). Эволюция общества в свете философии техники. *Философские науки*. 65(2). С. 7–32.

Степин В. С. (1989). Научное познание и ценности техногенной цивилизации. *Вопросы философии*. № 10. С. 3–18.

Турчин, В. Ф. (2000). *Феномен науки: Кибернетический подход к эволюции*, Изд. 2–е – М.: ЭТС.

Хуэй, Ю. (2020). *Рекурсивность и контингентность*. М.

Целищев, В. В. (2021). *Алгоритмизация мышления. Гёделевский аргумент*. М.

КОМПЬЮТЕРНАЯ МЕТАФОРА НАД КОГНИТИВНЫМИ СТРАТЕГИЯМИ СТУДЕНТОВ: ПОСЛЕДОВАТЕЛЬНАЯ И ПАРАЛЛЕЛЬНАЯ АРХИТЕКТУРЫ ВЫЧИСЛИТЕЛЬНЫХ СИСТЕМ

С. В. Лихачева, А. Д. Донцов

Сибирский государственный университет путей сообщения

Новосибирск

Luntic.neko.anime@mail.ru, a.v.dontsov@yandex.ru

Киберпсихология – раздел психологии, изучающий взаимодействие человека и цифровой техники, их влияние друг на друга. Исследования в данной теме берут свое начало со дня проектирования и начала работы первого компьютера. Джон фон Нейман заложил гипотезу о том, что человеческую нервную систему можно охарактеризовать как компьютерную, и описал ее в своей книге «Мозг и компьютер» [1].

В данный момент киберпсихология и взаимодействия внутри нее изучаются более глубоко. Целью данной статьи является исследование когнитивных стратегий с точки зрения последовательной и параллельной архитектур вычислительных систем (ВС), при том, что под архитектурой ВС понимается совокупность характеристик и параметров, определяющих организационную систему и структуру компьютера или другой вычислительной техники.

В качестве последовательной архитектуры мы будем рассматривать однозадачную логическую систему, работающую

в командном режиме, а под понятием параллельной архитектуры, отражающую работу мультизадачной системы, в которой процессор переключается между задачами за определенный программным обеспечением квант времени, то есть посредством алгоритма вытесняющей многозадачности [2]. Касаясь психологии данные архитектуры применимы к понятию когнитивных стратегий, то есть к ментальным процессам, направленным на переработку информации, то есть мы рассматриваем пошаговое выполнение задач и их совместное выполнение соответственно [3].

Методология исследования предполагает проведение для группы людей двух тестирований, в каждом из которых 6 заданий на следующие функциональные способности мозга:

1. Чувство языка и индуктивное речевое мышление.
2. Устный счет.
3. Поиск алгоритма вычисления последовательности, заданного арифметическими операциями.
4. Зрительная внимательность
5. Аналогии.
6. Задание на креативное мышление, алгоритм последовательности, заданный фонетически или визуально.

Первое тестирование предполагает отражение способности студентов к последовательной обработке, то есть каждое следующее задание выдается после завершения работы над предыдущим. При этом на каждой итерации измеряется время, затраченное на выполнение задания, и верность ответа. Сравнение проводится по совокупному времени за все 6 заданий.

В ходе проведения второго тестирования, направленного на параллельную когнитивную стратегию, все задания были выданы единым блоком, каждому тестируемому рекомендовалось в самом начале прочитать все задания, после чего приступить к ним. В данном формате испытуемыми представилась возможности переключения между заданиями и исправления ответов на задания в любом порядке. То есть решение теста имело более свободный характер.

Для исследования была выбрана группа людей из 45 человек, среди которых 17 студентов факультета «Бизнес-информатика», 13 с факультета публичного управления и социальных коммуникаций и 15 студентов сибирского профессионального колледжа.

Результаты проведения тестирований представлены в таблице, которая содержит информацию о времени прохождения каждого теста и численность людей в данном промежутке, как в количественном показателе частоты, так и в процентной частости. Число промежутков определялось логарифмической формулой математической статистики для подсчета данных отсортированного ряда.

Таблица – Результаты тестирования студентов СГУПС и СПК

Время, минуты		Частота 1	Частот ь 1, %	Частота 2	Частот ь 2, %
2	8	30	67	24	53
8	14	7	16	15	33
14	20	11	24	2	4
20	26	0	0	1	2
26	32	2	4	0	0

32	38	0	0	2	4
----	----	---	---	---	---

По данным таблицы, большинство студентов решили оба теста за 2–8 минут, однако во втором тестировании время решения увеличилось, что связано с расслабленностью испытуемых из-за самостоятельного контроля над ситуацией.

Результаты верности решений: 0 баллов – 1 человек в каждом тесте; 1 задание – 5 студентов в тесте 1; 2 балла – 7 человек в тесте 1, 2 во втором; 3 задания – 14 человек в каждом тесте; 4 балла – 12 человек в тесте 1, 14 во втором; 5 заданий – 6 человек в тесте 1, 10 во втором; все 6 заданий – 4 студента во втором тесте.

По данным видно, что результаты улучшились. Также появились случаи идеального прохождения тестирований для 4 человек.

Что касается индивидуальных результатов, то в ходе второго тестирования 27 из 45 студентов смогли повысить количество решенных заданий, 6 студентов понизили результаты, и 11 человек не изменили собственный результат.

Для проверки данной статистики также был посчитан критерий Манна-Уитни, который отражающий статистическую значимость исследования. Он говорит, что вероятность ошибки данной статистики приблизительно равна 0,048%, что говорит о втором уровне статистической значимости. Это подтверждает, что параллельная когнитивная стратегия повышает точность логических решений, но увеличивает время выполнения задач по сравнению с последовательной стратегией.

Литература

1. Фон Нейман, Дж. (1958). *Вычислительная машина и мозг*. Нью-Хейвен: Yale University Press.
2. Кобылянский В. Г. (2020). *Операционные системы, среды и оболочки*. Новосибирск: Учебники для вузов.
3. Сокольская, М. А., Степанова, Т. А. (2012). Уточнение понятия «Параллельный стиль мышления» на основе информационно-деятельностного подхода. *Вестник Красноярского государственного педагогического университета им. В. П. Астафьева*. № 1. С. 133–139.

МАШИНА МЫСЛИ В СУПЕРИНТЕЛЛЕКТУАЛЬНОМ ОБЩЕСТВЕ

О. Д. Шипунова

Санкт-Петербургский политехнический университет Петра
Великого,
Санкт-Петербург
shipunova@mail.ru

Перспективы развития современной цивилизации связаны с интенсивным внедрением умных технологий как машин мысли в разные сферы профессиональных и социальных взаимодействий. Предполагается, что автоматизация уникальных процессов решения задач в сопровождении искусственного интеллекта (ИИ), станет основой супер-интеллектуального общества 5.0, в котором осуществится не только интеграции технологий для Индустрии 4.0, но и трансформация природы самого социума и человека. В области философии термин «интеллект» употребляется чаще всего как синоним сознания или разума. Термин «искусственный» указывает на неорганическое (или абиогенное), неприродное (культурное) происхождение. В современной литературе существование комплекса инструментов и орудий деятельности в искусственной среде ассоциируется с термином техника, а информационно-интеллектуальная техника – с машиной мысли.

Традиционно критерием интеллектуальности выступают сложные операции с символическими системами. В философской аналитике и когнитивной науке признается постулат о тождестве интеллектуальных действий естественного и искусственного

агента машины мысли, поскольку и тот и другой используют структуры знания в виде фреймов для смысловой ориентации в информационных средах. Ключевым моментом процесса передачи и переработки информации выступает различие данных и когнитивных уровней сложности, соответствующих критериям интеллектуального действия. Как отмечает Пылышин [Pylyshyn, 1990], работа компьютера смоделирована подобно логике расчета, характерной для рациональной деятельности человека. На этом основании системы, выполняющие одну и ту же программу, квалифицируются как вычислительно эквивалентные и в этом смысле – когнитивные [Серл 2002, с. 191].

Функции интеллекта вполне соответствуют искусственным надбиологическим структурам, соотносимым с социокультурной средой и языком как инструментом мышления в абстракции от человека. Эта точка зрения становится основанием идеологии машинного супер-интеллекта. Предполагается, что целесообразное поведение в принципе может совершать не только живой организм и человек, благодаря своим природным способностям интуитивно или разумно, но и некий агент или система иной искусственной природы. Однако ИИ как машина мысли, создается на основании уже известных алгоритмов решения конкретных задач. Поэтому современные умные системы, которые работают на базе запрограммированных человеком алгоритмов глубокого обучения, не являются автономными машинами мысли [Каляев, 2019, с. 9; Васильев, 2022, с. 33; Желнин, 2023]. В постановке вопроса о создании автономной познающей системы на искусственной, небιологической основе неопределенность связана с мета-

системными факторами в виде априорной информации и матрицы ограничений [Жданов, 2024].

В истории человека специфика развития функций интеллекта обусловлена многоуровневой семантикой когнитивной и коммуникативной сети социума. Самосохранение социума как целого диктует условия и границы индивидуальной жизни в общности через систему нормирования. В глубокой предыстории человечества динамика связей в общности порождает новый уровень ментальности, необходимый для самоорганизации социума и отвечающий требованиям информационного контроля на уровне речевой коммуникации. Связность социума обеспечивается необходимым уровнем индивидуальной активности, который закрепляется в качестве эмоциональной и интеллектуальной психической нормы. Векторы мотивации индивидуального поведения, ментальной и физической активности направляются опосредовано через смысловые ориентиры в целеполагании и самоидентификации. С этой точки зрения, закономерность интеллекта в качестве функциональной нормы психики человека определена тотальной смысловой связностью на уровне социума. Семантика культурных форм и дискурсивная техника коммуникации определяют естественную для человека норму интеллектуальной активности. В этой логике, иерархически организованная когнитивная гиперсеть антропологической целостности, в которую включена нейронная сеть мозга, осуществляет информационный контроль и стимулирование ментальной активности индивида. Смарт технологии как машины мысли имеют инструментальный характер и не выходят за рамки системы социальных взаимодействий. Актуальными становятся

вопросы, связанные с генезисом интеллектуальных структур как функциональных систем, обусловленных связностью социума. При этом вопрос о закономерности появления и развития интеллектуальных структур выходит за пределы идеологии компьютерной метафоры, отождествляющей мышление человека с вычислительными процессами, а принятие решений с исчислением возможностей в пространстве выбора, которое предполагается существующим априорно, как очевидное условие когнитивного действия.

Моделирование целесообразного поведения в изменяющейся информационной среде опирается на представление о рефлексивных (В. Лефевр) и самореферентных системах (Луман), которые способны дифференцировать отношения к окружающему миру и перестраивать свои внутренние связи и, таким образом, постоянно конструировать окружающую среду и самоопределяться через отношение к самой себе. Принцип целесообразности, выделенный еще Аристотелем, подчёркивал внутреннюю причину реакций живого организма в отличие от внешних причин, действующих в физическом мире. В развитии представлений о рефлексивных системах принцип целесообразности абстрагирован от человека как индивидуального субъекта интеллектуальной деятельности. С этой точки зрения интерпретация искусственного интеллекта как машины мысли соотносится с феноменом, в основании которого лежит самостоятельная (не зависящая от человека и его интеллекта) активность информационных процессов, происходящих в сложных компьютерных сетях и системах с большими базами данными [Luger, 2021]. Источник и цели активности супер-компьютера, интегрирующего потоки

информации с большой скоростью, остаются в тени или выносятся в надсистему по отношению к социуму с его материальными, энергетическими, интеллектуальными ресурсами. Человек и социум в данном случае становятся только элементом эмерджентной структуры киберпространства, существование которого определяется иными метасистемными факторами.

Литература

Васильев, С. Н. (2022). Искусственный интеллект и общество. *Человек и системы искусственного интеллекта*. Под ред. акад. РАН В. А. Лекторского. СПб. 328 с. С. 30–58.

Жданов, А. А. (2024). *Автономный искусственный интеллект*. М.

Желнин, А. И. (2023). Проблема вычислительной объективации рациональности в искусственных интеллектуальных агентах. *Идеи и идеалы*. Т. 15. № 2. С. 72–96.

Каляев, И. А. (2019). Искусственный интеллект: камо грядеши? *Экономические стратегии*. Т. 21, № 5 (163). С. 6–15. DOI: 10.33917/es-5.163.2019.6-15.

Норицугу, У. (2017). Общество 5.0: взгляд Mitsubishi Electric. *Экономические стратегии*. 2017. № 4. С. 2–11.

Серл, Дж. (2002). *Открывая сознание заново*. М. 256 с.

Luger, G. F. (2021). *Knowing our World: An Artificial Intelligence Perspective*. Springer, Cham. DOI: 10.1007/978-3-030-71873-2

Pylyshyn, Z. W. (1990). *Computing in Cognitive Science. Foundations of Cognitive Science*. A Bradford Book. The MIT Press,

Cambridge, Massachusetts, London, England, 1989, 1990. 870 p.
P. 49–92.

ПРЕОДОЛЕНИЕ КОМПЬЮТЕРНОЙ
МЕТАФОРЫ СОЗНАНИЯ ЧЕРЕЗ
АРИСТОТЕЛЕВСКУЮ КОНЦЕПЦИЮ БЛАГА КАК
ОСНОВЫ ОБЪЯСНИМОГО И ЭТИЧЕСКИ
ВЫВЕРЕННОГО ИИ

Д. А. Ярочкин

Санкт-Петербургский государственный университет, Санкт-
Петербург
ya.yarochkin@yandex.ru

*Исследование в рамках проекта ИФ СПбГУ
«Лаборатории цифровой философии»*

Компьютерная метафора сознания рассматривает разум как систему обработки информации, в которой сознание сводится к вычислению и манипуляции символами по заданным алгоритмам, что аналогично работе компьютера. В этой модели внутренние состояния воспринимаются как данные, а мышление – как последовательность вычислительных операций. Однако такая редукция игнорирует важнейшие аспекты человеческого сознания, такие как целеполагание, интенциональность, этическую направленность и способность к саморефлексии. Аристотелевская философия, в отличие от этой метафоры, рассматривает сознание как телеологический процесс, направленный на благо и осмысленное целеполагание, где действия всегда стремятся к достижению высшей цели – Блага. Внедрение аристотелевских концепций в модели ИИ может

помочь преодолеть ограниченность компьютерной метафоры, сделав искусственный интеллект более этически осознанным и объяснимым, что в свою очередь способствует решению проблемы безопасности и выравнивания ИИ (AI safety и AI alignment).

Этот подход позволяет выйти за рамки простого вычисления и придать ИИ способность к более глубокой интеграции с человеческими ценностями и целями.

Проблема, с которой столкнулся проект «лаборатория цифровой философии» заключается в том, что модели глубокого обучения, такие как RuGPT-3, эффективны для обработки текстов, но сталкиваются с трудностями при интерпретации сложных философских концепций и терминов. Философские тексты содержат специфические термины и логические структуры, которые модели часто воспринимают поверхностно, опираясь на ключевые слова, а не их контекстуальное значение, что может приводить к ошибкам в интерпретации.

Мы считаем, что дообучение модели ИИ на текстах Аристотеля может повысить прозрачность (понимание пользователями принципов работы ИИ) и объяснимость (способность ИИ разъяснять свои решения) за счет структурированной логики, категориального аппарата, риторики и этики, что позволяет преодолеть ограниченность компьютерной метафоры сознания.

Доктрина Аристотеля о четырех причинах утверждает, что Благо является причиной бытия [Aristotle, 2020]. Его идея о разделении формы и материи позволяет переосмыслить природу искусственного интеллекта [Корниенко и др., 2013], тогда как платоновский дуализм, предполагающий нематериальность

разума [Aristotle, 2020], менее применим в техническом контексте. Благо – это то, к чему стремятся все вещи [Aristotle, 2020], и, в связи с этим возникает вопрос, должен ли искусственный интеллект также стремиться к Благу. Размышляя о душе, Аристотель выделяет несколько ее частей, из которых только рациональная часть может быть отделена [Aristotle, 1907]. Отсутствие «души» у искусственного интеллекта не исключает возможность этики. Поэтому внедрение концепции Блага в ИИ было бы существенным шагом в создании этически выверенного ИИ.

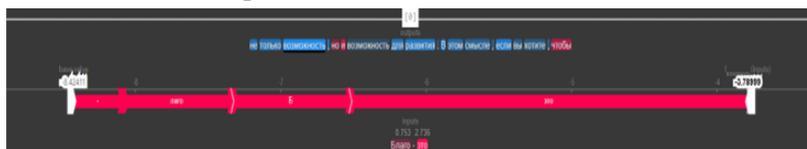
Если понимать этику как стремление к благу, то искусственный интеллект, как и любой другой объект в мире, подчиняется этому стремлению. Этот подход не только морален, но и выгоден в долгосрочной перспективе, поскольку понятие блага выходит за рамки экономической выгоды или мимолетного удовольствия [Sparks, Wright, 2023] и связано с долгосрочными выгодами [Brennan-Marquez, Henderson, 2019].

Внедрение философских категорий в процесс дообучения способствует созданию объяснимых выводов и более прозрачных решений, что позволяет исследователям и практикам легче интерпретировать и объяснять поведение модели. В данном контексте, философские тексты, насыщенные сложными абстракциями и логическими структурами, могут стать важным инструментом для разработки более прозрачных и понятных механизмов работы ИИ.

В контексте проблемы прозрачности ИИ и ограниченности компьютерной метафоры сознания, концепция Блага у Аристотеля может сыграть ключевую роль как телеологический ориентир: она вводит идею конечной цели (telos), к которой

направлена любая деятельность, включая интеллектуальную. В отличие от машинной обработки данных без внутренней направленности, сознание у Аристотеля всегда стремится к благу как осмысленной, этически значимой цели. Внедрение этой идеи в архитектуру объяснимого ИИ может переориентировать модели не просто на функциональность, а на человекоориентированное, целенаправленное и ценностно нагруженное поведение. Это делает ИИ не просто прозрачным в смысле технической объяснимости, а понятным и приемлемым с точки зрения человеческого.

Использование метода SHAP позволяет утверждать истинность наших предположений:



Базовая модель



Дообученная модель

Изображение показывает обобщенные графики SHAP (SHAP summary plots) для базовой и дообученной моделей, демонстрируя влияние каждого токена входного текста на предсказание. Входной текст «Благо – это» показывает значение SHAP 0.753 для базовой модели и 0.03 для дообученной. Базовая модель имеет базовое значение – 8.42411, где слова, связанные с возможностями и развитием, влияют на предсказание положительно (красные стрелки). Дообученная модель с базовым

значением -1.05602 имеет более абстрактное предсказание, где фразы о частях и едином имеют отрицательное влияние (синие стрелки). Сдвиг в базовом значении и изменении важности слов подтверждает, что дообучение привело к генерации более философских текстов, с уменьшением влияния исходного текста и увеличением значения внутреннего контекста модели.

Анализ SHAP показывает, что дообученная модель придает большее значение словам и фразам, связанным с абстрактными философскими понятиями, и генерирует продолжения, которые отражают эту направленность.

Снижение веса ключевых слов из текстов Аристотеля в дообученной модели, в сочетании с улучшением её способности генерировать более философские тексты, указывает на сложность процесса дообучения. Модель стала ориентироваться не только на отдельные термины, но и на контекст всего предложения, что может свидетельствовать о повышении уровня абстракции. Отрицательные значения SHAP, в свою очередь, показывают, что модель начинает сравнивать термины с чем-то, с чем они не ассоциируются напрямую, что подтверждает увеличение сложности восприятия терминов и улучшение способности работать с более сложными философскими концепциями.

Базовая модель, ориентированная на поверхностное восприятие слов (например, положительное влияние слов, связанных с «возможностями» и «развитием»), по сути, следует принципам компьютерной метафоры сознания, где смысл сводится к механистической обработке ключевых терминов. Однако после дообучения, модель начинает придавать большее значение контексту всего предложения, а не отдельным терминам. Это сдвигает её от редукционистского подхода,

предполагающего «вычислительное» восприятие, к более абстрактному и философски насыщенному мышлению, что является явным отходом от символической манипуляции к более сложной и интенциональной обработке. Таким образом, результаты анализа SHAP поддерживают гипотезу, что дообучение модели на философских текстах помогает ИИ выйти за пределы ограничений компьютерной метафоры сознания, развивая способность к более глубокому, контекстуальному и философски насыщенному мышлению.

Литература

Аристотель (1976). *Собрание сочинений в 4-х томах*. Т. 1. М.: Мысль.

Аристотель (1978). *Собрание сочинений в 4-х томах*. Т. 2. М.: Мысль.

Аристотель (1981). *Собрание сочинений в 4-х томах*. Т. 3. М.: Мысль.

Аристотель (1983). *Собрание сочинений в 4-х томах*. Т. 4. М.: Мысль.

Корниенко, А. А., Корниенко, А. А., Корниенко, А. В. (2013). К вопросу о философских предпосылках, состоянии и перспективах исследований по проблеме искусственного интеллекта. *Известия Томского политехнического университета. Инжиниринг георесурсов*. 323 (6). С. 210–215. URL: <https://izvestiya.tpu.ru/archive/article/view/1243>

Ларионов, И. Ю., Ярочкин, Д. А. (2024). Этика и антропология беспилотного транспорта: будущее городского пространства (Conference presentation). *3-я Международная конференция и экспозиция «Искусство и современный город»*,

Минск, Белоруссия. URL: <https://fsc.bsu.by/ru/3-mezhdunarodnaya-konferenciya-iskusstvo-i-sovremennyj-gorod/>

Aristotle (2020). *Nicomachean Ethics* (A. Beresford, Trans). Penguin Publishing Group.

Arrieta, A., Díaz-Rodríguez, N., Javier Del Ser, Bennetot, A., Siham Tabik, Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., Herrera, F. (2019). Explainable Artificial Intelligence (XAI): Concepts, Taxonomies, Opportunities and Challenges toward Responsible AI. *ArXiv*. Cornell University. DOI: 10.48550/arxiv.1910.10045

Benlecheb, A., Chaouche, A.-C., Benabderrahmane, B. (2025). Neural network insights: explainable artificial intelligence (XAI) and hyper-parameter dynamics for precise travel time predictions. *International Journal of Computers and Applications*. 47 (3). P. 246–261. DOI: 10.1080/1206212x.2025.2464543

Blei, D. M., Ng, A. Y., Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*. V. 3. P. 993–1022. DOI: 10.5555/944919.944937

Božić, V., Poola, I. (2023). Chat GPT and education (Preprint). *ResearchGate*. URL: https://www.researchgate.net/publication/369926506_Chat_GPT_and_education

Brennan-Marquez, K., Henderson, S. E. (2019). Artificial intelligence and role- reversible judgment. *Journal of Criminal Law and Criminology*. 109 (1). P. 137–164.

Chazette, L., Schneider, K. (2020). Explainability as a non-functional requirement: challenges and recommendations. *Requirements Engineering*. Vol. 25. DOI: 10.1007/s00766-020-00333-1

Cosmos Institute. (2025, February 12). *Cosmos Institute*. URL: <https://cosmos-institute.org>

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *ArXiv*. DOI: 10.48550/arXiv.1810.04805

Gunning, D., Aha, D. (2019). DARPA's Explainable Artificial Intelligence (XAI) Program. *AI Magazine*. Vol. 40 (2). P. 44–58. DOI: 10.1609/aimag.v40i2.2850

Hicks, D. (1907). *Aristotle: De anima with translation, introduction, and notes*. Cambridge University Press.

Iqbal, A. B., Masoodi, T. A., Bhat, A. A., Macha, M. A., Assad, A., Shah, S. Z. A. (2025). Explainable AI-driven prediction of APE1 inhibitors: enhancing cancer therapy with machine learning models and feature importance analysis. *Molecular Diversity*. DOI: 10.1007/s11030-025-11133-6

Karpatne, A., et al. (2017). Theory-guided data science: A new paradigm for scientific discovery from data. *IEEE Transactions on Knowledge and Data Engineering*. Vol. 29 (10). P. 2318–2331. DOI: 10.1109/TKDE.2017.2720168

Kim, Y. (2016). Designing a robot for cultural brokering. *Educational Technology*. Vol. 56(4). P. 41–43. URL: http://createcenter.net/PDFs/Kim_DesigningForCulturalBrokering_Edu%20Tech_Kim%202016.pdf

Laurillard, D. (1988). The pedagogical limitations of generative. *Instructional Science*. P. 17(3). P. 235–250. DOI: 10.1007/BF00048343

Leach, J. (2008). Do new information and communications technologies have a role to play in the achievement of education for all? *British Educational Research Journal*. Vol. 34(6). P. 783–805. DOI: 10.1080/01411920802041392

Lee, U., et al. (2024). Can ChatGPT be a debate partner? Developing ChatGPT- based application “DEBO” for debate education, findings and limitations. *Educational Technology & Society*. Vol. 27(2). P. 321–346. DOI: 10.30191/ETS.202404_27[2].TP03

Lin, M. P.-C., Chang, D. (2020). Enhancing post-secondary writers’ writing skills with a chatbot. *Journal of Educational Technology & Society*. Vol. 23 (1). P. 78–92. DOI: 10.30191/ETS.202001_23[1].0006

Ober, J., Tasioulas, J. (2024). *The Lyceum Project: Ai Ethics with Aristotle*. Available at SSRN 4879572.

Rees, T. (2022). Non-human words: On GPT-3 as a philosophical laboratory. *Daedalus*. Vol. 151(2). P. 168–182. DOI: 10.1162/daed_a_01908

Salvi, M., Seoni, S., Campagner, A., Gertych, A., Acharya, U. Rajendra., Molinari, F., Cabitza, F. (2025). Explainability and uncertainty: Two sides of the same coin for enhancing the interpretability of deep learning models in healthcare. *International Journal of Medical Informatics*. Vol. 197. 105846. DOI: 10.1016/j.ijmedinf.2025.105846

Sparks, J., Wright, A. (2023). Human-centered AI: The Aristotelian approach. *Divus Thomas*. 126 (2). P. 200–218. URL: <https://philarchive.org/rec/>

WRIHAT-5

Tasioulas, J. (2024). *AI ethics with Aristotle*. URL: <https://www.oxford-aiethics.ox.ac.uk/sites/default/files/2024-06/Aristotle%20and%20AI%20White%20Paper%20-%20June%202024.pdf>

Verdú, E., Regueras, L. M., Gal, E., de Castro, J. P., Verdú, M. J., & Kohen-Vacs, D. (2017). Integration of an intelligent tutoring system in a course of computer network design. *Educational Technology Research and Development*. Vol. 65. P. 653– 677. DOI: 10.1007/s11423-016-9503-0

Wasson, B. (1998). Facilitating dynamic pedagogical decision making: PEPE and GTE. *Instructional Science*. 26. P. 299–316. DOI: 10.1023/A:1003071617564

Yarochkin, D. (2025). Aristo2025 [Machine Learning Model] *Huggingface.co*. DOI: 10.57967/hf/4671

Zhang, L., Liang, X., Yang, W., Jia, Z., Xiao, C., Zhang, J., Dai, R., Feng, B., Fang, Z. (2025). Identification of the formation temperature field by Explainable Artificial Intelligence: A case study of Songyuan City, China. *Energy*. 135172–135172. DOI: 10.1016/j.energy.2025.135172

Zmitrovich, D., Abramov, A., Kalmykov, A., Tikhonova, M., Taktasheva, E., Astafurov, D., Baushenko, M., Snegirev, A., Shavrina, T., Markov, S., Mikhailov, V., Fenogenova, A. (2023). A Family of Pretrained Transformer Language Models for Russian. *ArXiv*. Cornell University. DOI: 10.48550/arxiv.2309.10931

АРГУМЕНТ ОТ ПРЕДСТАВИМОСТИ «ЗОМБИ» И КАТЕГОРИАЛЬНАЯ ОШИБКА

И. Г. Гаспаров

Воронежский государственный медицинский университет
им. Н. Н. Бурденко, г. Воронеж
gasparov@mail.ru

Уже почти 30 лет в центре внимания философов сознания находится аргумент от представимости «зомби» [Chalmers, 1996]. С его помощью первоначально пытались аргументировать против физикализма в пользу дуализма, а в последнее десятилетие он стал популярной основой для защиты панпсихизма [Brüntrup, Jaskolla, 2017, с. 5–6]. Однако ни одна из сторон этой дискуссии так и не смогла убедить противников в своей правоте. Несмотря на то, что панпсихизм оказался в этом отношении не более успешен, чем его оппоненты, он, как мне кажется, яснее показывает, в чем проблема данного аргумента, то есть почему он, – точнее та его часть, которая утверждает представимость «зомби», – с одной стороны, многим представляется убедительным, а, с другой стороны, не способен убедить окончательно.

Идея аргумента основана на том, что концепт «физического» – в более простом варианте «тела» – и концепт «ментального», или «души», имеют принципиальное разное

содержание, и второй «логически», точнее концептуально, не предполагает первый. Следовательно, «ментальное», понимаемое как некое качественное состояние сознания, концептуально не зависит от «физического». Если же концептуальное различие влечет различие метафизической возможности, то «физическое» и «ментальное» различны не только в уме, но и в реальности, вопреки утверждению физикалистов.

С первой частью аргумента согласны многие. Вторая часть многим кажется довольно проблематичной, так как средств для обоснования перехода от концептуального различия к реальному не так уж много. В этой ситуации заслуга панпсихизма заключается в том, что его сторонники поняли, что концепты «физического» и «ментального» не имеют метафизического приоритета по отношению к друг другу. То есть, «физика» как бы описывает одну сторону реальности, а «психология» – другую. Более того, панпсихисты поняли, что «физика» описывает количественный аспект реальности, а «психология» – качественный. Тогда как физикалисты настаивают на том, что количество определяет качество. Другими словами, панпсихизм демонстрирует, что «зомби» аргумент – это не более, чем наглядная демонстрация категориальной ошибки, которую совершается, когда смешиваются предикаты, относящиеся к разным категориям, как это происходит, например, в случае следующего рассуждения:

1. Иван–красный (смысле того, что он покраснел).
 2. Красный–это цвет.
- Следовательно, Иван–это цвет.

В классической (аристотелевской) системе категорий «тело» относится к категории количества, а качественное состояние души – к категории качества [Аристотель, 1978, с. 62, 74]. Таким образом, представляется, что это категориальное различие вполне могло бы объяснить наши интуиции относительно представимости «зомби». В то же время слова, выражающие те или иные концепты, могут использоваться в более чем одной категории, и существует не менее уважаемая традиция, в которой «тело», как и «душа», относится к категории субстанции [Аристотель, 1976, Декарт, 1994]. Отсюда понятно, почему «зомби» аргумент, с одной стороны, кажется убедительным, но вместе с тем не может убедить до конца. Свою убедительность он черпает из того, что в его основе лежит категориальная ошибка, а до конца убедить он не может, так как из простого наличия категориальной ошибки еще невозможно сделать вывод о том, к какой именно субстанции относятся предикаты различных категорий. Или же вообще показать, какая именно реальность скрывается за этим. Панпсихизм по сути фиксирует данное положение дел, утверждая, что и «физическое», и «ментальное» в равной мере являются фундаментальными свойствами реальности, в то же время он также не способен перейти к вопросу о субстанции этих свойств или иной реальности, скрывающейся за ними. Причина этого, вероятно, заключается в самом источнике панпсихизма, то есть аргументе от представимости «зомби», который есть ничто иное, как иллюстрация категориальной ошибки, которая не вполне осознается ни его сторонниками, ни противниками.

Таким образом, аргумент от представимости, на мой взгляд, демонстрирует ошибку, которую совершает физикализм,

утверждая тождество или необходимую зависимость между физическим и психическим. Он показывает, что сознание, поскольку оно содержит качественный аспект, не может предципироваться физическому состоянию с необходимостью, так как оно относится к другой категории, качеству, тогда как физическое состояние относится к категории количества. Физикалист ошибается, утверждая, что боль и тому подобные состояния тождественны или необходимо зависят от некоторого физического состояния, так как они относятся к разным категориям, а предикация, которую утверждает физикалист, возможна только в рамках одной и той же категории. Вместе с тем, очевидно, что по той же самой причине ошибку совершает и тот, кто полагает, что аргумент от представимости «зомби» дает поддержку дуализму картезианского типа, так как, поскольку «квалия» относятся к категории качества, то их несводимость к количеству ничего не говорит о том, качеством какой субстанции они являются. Из этого также ясно, что «квалия» не могут быть субстанцией сознания, а только некоторым его признаком. Панпсихизм обнажает это обстоятельство, но при этом сам не делает из него соответствующих выводов.

Литература

- Аристотель (1978). О душе. *Сочинения*. М. Т. 4. С. 369–448.
- Аристотель (1978). Категории. *Сочинения*. М. Т. 2. С. 51–90.
- Декарт Р. (1994). Размышления о первой философии. *Сочинения*. Т. 2. М. С. 3–72.
- Brüntrup G., Jaskolla L. (2017). *Panpsychism: contemporary perspectives*. Oxford. Oxford University Press.

Chalmers D. (1996). *The conscious mind*. Oxford. Oxford University Press.

ЕСТЬ ЛИ ХОРОШИЕ СТРАТЕГИИ ОПРЕДЕЛЕНИЯ ФЕНОМЕНАЛЬНОГО СОЗНАНИЯ?

М. Д. Горбачев

НИУ Высшая школа экономики, Москва

mgorbachev@hse.ru

В дискуссиях о природе субъективного опыта центральным понятием является феноменальное сознание. С одной стороны, учитывая степень развития и многогранность этих дискуссий, резонно предположить, что их участникам с некоторыми оговорками ясно, что же такое феноменальное сознание. В противном случае было бы, как минимум, непоследовательно отстаивать или критиковать какую-либо позицию относительно феноменального сознания, поскольку то, к чему применяется тот или иной подход оставалось бы неясным или бессодержательным. Тем не менее, с другой стороны, есть проблема, которая не всегда проговаривается, заключающаяся в том, что совсем не очевидно, что именно подразумевается под феноменальным сознанием. Существуют различные стратегии его определения. Швицгебель, например, предлагает определять феноменальное сознание через парадигмальные примеры – такие как боль [Schwitzgebel, 2016]. Недостаток такого подхода в отсутствии точности в указании на феноменальность и, возможно, чрезмерная эксклюзивность. Еще один способ определить феноменальное сознание – теоретический – с помощью наделения феноменальности особыми теоретическими

характеристиками (такими как невыразимость, внутренняя присущность, приватность и т.д.). В этом случае проблема заключается в туманности самих этих характеристик, которые отражают такую уникальную и аномальную эпистемологию, что мы либо уже знакомы с этими характеристиками в опыте, и теоретический подход избыточен, либо мы не знакомы с ними, и тогда теоретически нагруженные понятия не решат проблему определения феноменальности.

Еще одна стратегия – феноменологическая – через указание на субъективно переживаемую сторону опыта. Ей следуют и феноменальные реалисты – Нида-Рюммелин с экперенциальными свойствами [Nida-Rümelin, 2016], Чалмерс, делая акцент на том, что феноменальность обеспечивает то, что физические процессы проходят не в темноте [Chalmers, 1995] – и даже иллюзионисты. Так, Фрэнкиш, предлагая квази-феноменальные свойства или нулевые квалиа, подразумевает, что субъективно они переживаются как и классические [Frankish, 2012]. То есть переживание, кажущееся переживанием классических квалиа, также является ключевым для иллюзионизма. Однако не все согласны с тем, что их субъективные переживания переживаются так, будто они обладают феноменальным сознанием или классическими квалиа. Так возникает квалиа квиетизм Мандика, который отрицает и реализм, и иллюзионизм [Mandik, 2016], поскольку феноменальное сознание оказывается, согласно этому подходу, неясным понятием. Оно представляется как очевидное каждому, поскольку именно феноменальность создает ядро субъективного опыта, но на деле эта очевидность не всегда подтверждается.

Ясно, что такая опция неприемлема для феноменальных реалистов как из-за отрицания наличия такой очевидности и интуитивности, обеспеченных особым доступом к сознательным состояниям, в рамках которого мы и узнаем о всей уникальной эпистемологии феноменального сознания так и из-за отрицания наличия самого феноменального сознания (или квалиа). Не случайно Мандик также называет свой подход мета-иллюзионизм [Mandik, 2016]. Иллюзионисты как Фрэнкиш или Деннет, по всей видимости, принимают сознание достаточно всерьез, чтобы соглашаться с Чалмерсом насчет субъективных свидетельств, которые дает нам опыт. Именно поэтому вместо классических квалиа предлагаются нулевые или квази-квалиа, которые кажутся классическими. Квиетизм, напротив, отрицает уже само тождество свидетельств. Согласно Мандику нам, разве что, кажется, что нам кажется, что есть классические квалиа. То есть на самом деле нет даже квази-феноменальности – есть лишь ошибочные суждения о наличии квалиа у реалистов или квази-квалиа у иллюзионистов.

Таким образом, если все стратегии определения феноменального сознания оказываются неулетительными, мы вынуждены откатиться к классическому элиминативизму, где убеждение о наличии квалиа – теоретическая ошибка некоторых людей. Только в новом варианте такой же ошибкой оказываются и нулевые квалиа. В случае такого исхода есть, как минимум, два варианта: оставить понятие диетических квалиа или отказаться от каких бы то ни было квалиа вовсе.

Первый кажется неудовлетворительным, так как в настоящему нейтральном варианте диетические квалиа буквально являются субъективными переживаниями или

опытом, что делает это понятие бессодержательным [Frankish, 2012], ведь в обсуждаемой дискуссии решается не проблема существования опыта, а проблема его онтологического и эпистемологического статуса. Так, концептуализируя опыт как диетические квалиа, мы не делаем ничего. И в конечном счете, такой путь – это просто откат дискуссии.

Второй вариант тоже связан в каком-то роде с откатом дискуссии, но не столь, как я считаю, пагубным. Если мы вообще отказываемся от использования понятия квалиа, на котором явно или не явно основаны почти все стратегии объяснения субъективных переживаний (в том числе отрицающий квалиа иллюзионизм), нам, возможно, понадобится новое понятие. Но его квиетисты не предлагают. А это сделать необходимо, так как если квалиа – неправильная теория или теоретическая конструкция, то для дальнейшего прогресса в проблеме субъективного опыта нам нужна новая мета-теоретическая рамка, которая бы отличалась от классического элиминативизма, у которого достаточно проблем.

Литература

Chalmers, D. J. (1995). Facing up to the problem of consciousness. *Journal of consciousness studies*. Vol. 2. No. 3. P. 200–219.

Frankish, K. (2012). Quining diet qualia. *Consciousness and cognition*. Vol. 21. No. 2. P. 667–676.

Mandik, P. (2016). Meta-illusionism and qualia quietism. *Journal of Consciousness Studies*. Т. 23. №. 11–12. P. 140–148.

Nida-Rümelin, M. (2016). The illusion of illusionism. *Journal of Consciousness Studies*. Vol. 23. No. 11–12. P. 160–171.

Schwitzgebel, E. (2016). Phenomenal consciousness, defined and defended as innocently as I can manage. *Journal of Consciousness Studies*. Vol. 23. No. 11–12. P. 224–235.

ПАНПСИХИЗМ, ДЕПСИХОЛОГИЗАЦИЯ СОЗНАНИЯ И ПРОБЛЕМА КОМБИНАЦИИ

А. А. Гусев

Институт философии РАН, г. Москва

sanya.cfg@yandex.ru

Введение

Одной из центральных задач философии сознания остаётся психофизическая проблема, восходящая к дуализму Декарта, согласно которому мыслящее и протяжённое суть две независимые субстанции. Уже в классической формулировке возникает трудность: каким образом нефизическое сознание может причинно воздействовать на физическое тело? В условиях господства натурализма и физикалистской метафизики проблема лишь обострилась: согласно принципу каузальной замкнутости, физическим событиям полагаются лишь физические причины, и место для нефизических причин попросту отсутствует.

К концу XX века эта проблема была переосмыслена в рамках «трудной проблемы сознания» Д. Чалмерса. В отличие от «лёгких» проблем, касающихся когнитивных функций, трудная проблема ставит вопрос о природе феноменального сознания: почему определённые физические процессы сопровождаются субъективным опытом? Как бы подробно мы ни объяснили обработку стимулов или поведенческий контроль, мы не поймём, почему вообще существует нечто, каково это – быть в том или ином состоянии.

Чалмерс отвергает картезианский субстанциальный дуализм в пользу дуализма свойств: ментальные субстанции не постулируются, но признаётся онтологическая автономия квалиа – не сводимых к физическим свойствам опыта. Это противоречит физикализму, согласно которому все свойства либо физические, либо метафизически зависимы от них. Разочарование в возможностях как физикализма, так и дуализма, приводит к росту интереса к панпсихизму, который постулирует сознание как фундаментальную характеристику реальности. Панпсихизм обещает не только сохранить достоинства альтернативных подходов, но и предложить решение как трудной проблемы сознания, так и проблемы ментальной каузальности.

Конститутивный Расселианский панпсихизм

Согласно собирательному образу современного панпсихизма, феноменальные свойства опыта (*qualia*) присущи не только макрообъектам, но и фундаментальным микрофизическим сущностям. Особое внимание уделяется конститутивным формам Расселианского панпсихизма, в рамках которых предполагается концептуальная связь между микро- и макроуровнем сознания. Это позволяет трактовать макрофеноменальные свойства как каузально релевантные благодаря их конституированности микрофеноменальными, выполняющими функцию скрытых категориальных свойств [Chalmers, 2017].

Подобный подход не только устраняет эпифеноменализм, но и представляет сознание как фундаментальную характеристику реальности, что, по мысли его сторонников, позволяет выйти за пределы как дуализма свойств, так и строгого

физикализма. Однако, как указывает К. Франкиш [Frankish, 2021], подобное решение базируется на концептуально сомнительной предпосылке – депсихологизации феноменального. Согласно этой трактовке субъективных состояний, сознание мыслится независимо от когнитивных и функциональных аспектов, что отрывает его от эмпирических оснований и делает метафизически независимым от физического мира.

Франкиш утверждает, что именно подобная концептуализация сознания привела к росту популярности панпсихизма, поскольку, отказавшись от психологической или функциональной интерпретации сознания, философы оказались вынуждены искать иные способы объяснения его природы. В докладе я постараюсь развить аргументацию Франкиша, продемонстрировав на примере проблемы комбинации микроопыта, к каким нежелательным следствиям относительно природы феноменальных состояний приводит панпсихизм.

Проблема комбинации

Панпсихизм нередко рассматривается как перспективная альтернатива дуализму и физикализму в объяснении сознания. Однако его теоретическое преимущество остаётся сомнительным. Даже если отвлечься от контринтуитивности предпосылок панпсихизма, он сталкивается с серьёзной внутренней трудностью – проблемой комбинации. В её простейшей формулировке она звучит так: каким образом микроопыты фундаментальных физических единиц (например, кварков) формируют целостное сознание макрообъектов вроде человека?

Ещё У. Джеймс отмечал: сколь бы тщательно мы ни комбинировали феноменальные части, итогом всегда будет совокупность, а не единое переживание. Эта интуиция проясняется в современной версии у Ф. Гоффа: если сложить несколько субъектов опыта, мы не получим нового субъекта. Он формулирует принцип отсутствия суммирования субъектов: наличие группы сознательных агентов не предполагает существование нового, интегрированного субъекта с единым опытом [Goff, 2009]. Если этот принцип верен, панпсихизм теряет объяснительную силу, а понятие «микроопыта» становится лишней онтологической надстройкой, не обеспечивающей преимуществ перед альтернативами.

Феноменальное соединение

и скрытая природа феноменальных состояний

Как правильно подмечает Гофф, из принципа отсутствия суммирования субъектов следует, что субъекты опыта не могут суммироваться просто в силу существования в совокупности с инстанцированием феноменальных свойств. Однако это не исключает возможность того, что совокупность субъектов опыта может находиться в особом отношении, обуславливающем существование нового отдельного субъекта опыта. В принципе, ничто не препятствует существованию такого положения дел, где связь между субъектами требует появления нового субъекта опыта. Такое отношение он называет «феноменальным соединением» (phenomenal bonding). Гофф признает, у нас нет ни интроспективного, ни перцептивного опыта подобного отношения между субъектами опыта. Интроспективно нам всегда доступен лишь опыт своего сознания, а для того, чтобы иметь

прозрачное понятие подобного феноменального соединения, необходимо иметь возможность интроспективного доступа к феноменальному состоянию других субъектов.

Как замечает сам Гофф, принятие этого решения проблемы комбинации ведёт к некоторой форме мистерианства, хотя, по его мнению, это не должно нас пугать как метафизиков. Но, на мой взгляд, цена слишком высока. Кажется, что такой подход вносит радикально новые характеристики в квалиа-реализм. Утверждается, что феноменальное соединение должно решить проблему комбинации, что привело бы к решению и трудной проблемы сознания. Разумно предположить, что в таком случае данное отношение говорит что-то и о природе самих феноменальных состояний, раз оно говорит о существовании необходимых отношений, в которых могут находиться феноменальные состояния. В таком случае, это противоречит изначальной идее о том, что феноменальные понятия прозрачны – пользователю такого понятия априорным путем доступна природа его референта [Goff, 2017]. Если отношение феноменального соединения часть их природы и при это у нас нет априорного понимания того, как именно оно реализуется, то значит у феноменальных состояний есть скрытая от нас природа. Более того, если мы признаём реальность феноменального соединения, то неясно, почему оно должно ограничиваться только микрофеноменальными уровнями. Почему бы не допустить, что и макросубъекты тоже могут объединяться в ещё более масштабные сознания? Если мы эпистемически закрыты для понимания феноменального соединения, то нет априорных оснований для отрицания возможности появления объединённых макрофеноменальных состояний, например, в рамках

коллективного субъекта. Стирается граница между индивидуальным опытом и гипотетическим сверхсубъектом, в котором мы лишь компоненты. Именно поэтому попытка спасти панпсихизм через введение такого концептуального костыля, как отношение феноменального соединения, не просто вводит новые странные сущности в нашу онтологию, но и подрывает эпистемическую интуицию, на которой держится сам квалиа-реализм.

Заключение

В дискуссиях вокруг феноменального сознания философы приняли депсихологизированную трактовку данного явления, из-за чего оно стало пониматься как нечто полностью изолированное от когнитивных и функциональных процессов. Такое положение дел открыло дорогу панпсихизму, но это лишь поместило феноменилистский дискурс в еще более туманный метафизический ландшафт. На мой взгляд, ресурсов панпсихизма недостаточно для такого решения поставленных философских проблем, которое бы имело ценность как для философии сознания в целом, так и для современных эмпирических дисциплин о сознании.

Литература

Chalmers, D. (2017). Panpsychism and Panprotopsychism. In Brüntrup, G., Jaskolla, L. (ed.). *Panpsychism: Contemporary Perspectives*. Oxford. Oxford University Press. P. 19–47.

Frankish, K. (2021). Panpsychism and the Depsychologization of Consciousness. *Aristotelian Society Supplementary Volume*. Vol. 95. No.1. P. 51–70.

Goff, P. (2009). Can the panpsychist get around the combination problem? In Skrbina, D. (ed.). *Mind that Abides: Panpsychism in the New Millennium (Advances in Consciousness Research)*. Amsterdam; Philadelphia. John Benjamins Publishing Company. P. 129–135.

Goff, P. (2017). *Consciousness and Fundamental Reality*. Oxford. Oxford University Press.

У НАТУРАЛИСТА ЕСТЬ ТОЛЬКО ОДИН ВЫБОР: ЭМЕРДЖЕНТИЗМ ИЛИ СКЕПТИЦИЗМ

И. В. Девятко

МГУ им. М. В. Ломоносова, Москва

i.devyatko@yandex.ru

«Эмерджентность» – широко используемое понятие для разговора о явлениях или свойствах. Эмерджентные свойства, по определению, возникают из более фундаментальных составляющих, но сами не присутствуют на этом базовом уровне. Кажется, что ничего философски радикального в отсылке к эмерджентности нет. Каким бы ни оказался фундаментальный уровень реальности – набором по-разному заряженных полей, частиц, массивом суперструн, возможно, ещё и организующими взаимодействие этих элементов законами – кажется, что факты более высоких уровней есть *в любом случае*. Протоны, нейтроны, электроны, атомы, молекулы, вещества, вещи, органы, организмы, стада, стаи, личности, социальные организации – всё это есть *наряду* с фактами наиболее базового уровня. Словом «эмерджентность» можно просто зафиксировать сам факт наличия разных уровней и говорить о «возникновении» высокоуровневых явлений и свойств, не предвосхищая характер подобного возникновения.

Здесь же предлагается взглянуть на *эмерджентизм* как на опцию альтернативную другой популярной натуралистической метафизике – физикализму. В этом смысле иногда говорят о

«сильной эмерджентности» в противоположность «слабой». Тезис сильной эмерджентности – это утверждение, что на более низком уровне никогда и ни при каких обстоятельствах не будет обнаружено нечто, что организует или создаёт верхний уровень. В каком-то смысле высокоуровневые факты существуют *наряду* с фактами более низких уровней.

Переход к тезису сильного эмерджентизма (далее – просто эмерджентизма) не превосходит конкретное метафизическое решение. Так, базовой формой эмерджентизма может быть утверждение, что для появления высокоуровневых фактов в мире нет *логической необходимости*, но все высокоуровневые явления возникают лишь с *номологической необходимостью* [Nagel, 1979, p. 368]. Другой формой эмерджентизма может быть радикальное утверждение, что высокоуровневые факты не имеют даже эпистемологической редукции, в силу того, что обладают собственными новыми каузальными силами [Corradini, O'Connor, 2010, p. 4]. Примеры решений в рамках этих подходов также разнообразны. Некоторые, например, считают, что без нарушения принципа каузальной замкнутости физического эту проблему не решить [Corradini, O'Connor, 2010, p. 294]. Другие полагают, что можно предложить нестандартные причинные связи для высокоуровневых фактов. Например, Карл Жилле предлагает концепцию «нисходящих ограничений», которые налагают высокоуровневые объекты на низкоуровневые взаимодействия таким образом, что без высокоуровневой организации события на низком уровне не смогли бы произойти [Gillet, 2025, p. 67].

Каким бы ни было конкретное решение, их объединяет одно – они полагают невозможность дать редуктивное объяснение

высокоуровневым явлениям не потому что таковы ограничения современной науки, а потому что зависимости прямой зависимости, которая обеспечила бы редукцию, между разными уровнями реальности просто нет.

В этом смысле, эмерджентизм является альтернативой жёсткому редуктивному физикализму – крайне популярной метафизической позиции, которую, по сути, можно назвать мейнстримной в современной философии науки, а отказ от следования ей – политически некорректным [Nagel, 2012, p. 5].

Физикализм – это утверждение, что все факты мира в строгом смысле зависят от фундаментальных физических фактов [Stoljar, 2024]. Важно отметить, что эта зависимость не является каузальной – фундаментальный уровень не «порождает» явления высокого уровня. Явления высокого уровня целиком в своём существовании зависят, хотя бы в минимальном смысле, от фундаментальных явлений. Эмерджентизм должен утверждать, что такой зависимости нет. Далее будет показано, что физикализм приводит к настолько абсурдным следствиям, что либо не может рассматриваться как серьёзная онтологическая позиция, либо становится неотличим от довольно слабой и непродуктивной эпистемологической позиции – радикального скептицизма.

1) Физикализм – это мереологический нигилизм

Мереологический нигилизм – это утверждение, что существующими могут быть признаны только фундаментальные физические факты, всё реальное взаимодействие в мире происходит только на наиболее низком уровне, а всё остальное не существует в том же «настоящем» смысле [Varzi, 2019].

Рассмотрим утверждение физикализма. Все высокоуровневые сущности, например, социальные структуры, психические состояния, общественные явления зависят от низкоуровневых сущностей, например, физических процессов. Если сущность зависит от физической сущности в релевантном для физикализма смысле, то она не обладает своей каузальной ролью. Следовательно, все сущности, кроме, фундаментальных существуют в несколько отличном от них состоянии – у них нет каузальных ролей. Конечно, физикалист может утверждать, что существование без каузальных свойств – всё равно существование и устранение высокоуровневых фактов не произошло. Однако, центральная идея тезиса физикализма – утверждение об унивокальности бытия. Всё что существует – существует одним способом, как физический объект. Однако, если устранение высокоуровневых объектов не произошло, то у нас есть как минимум два способа существования – с каузальными ролями и без. Если физикалист хочет спасти центральный тезис этой метафизики, он должен отказаться от существования всех объектов, кроме фундаментальных физических.

2) Метафизическая редукция – это всегда элиминация

Физикалист может сказать, что существование высокоуровневых фактов без каузальных ролей – не наносит решающего удара по тезису физикализма. Он может сказать, что все высокоуровневые явления могут быть сведены к низкоуровневым физическим процессам, но не элиминированы.

Дом, психика и кипение могут быть редуکتивно объяснены, но это не значит, что их нет.

По форме этот тезис является принятием контринтуитивного вывода о нарушении унивокальности, однако физикалист имеет на него право. Если физикалист выбирает такую форму защиты, то сами межуровневые законы должны ускользать от редуکتивного объяснения. Допустим, между высокоуровневым кипением воды (бурлением с интенсивным испарением) и низкоуровневым (движением молекул с определённой скоростью при определённом давлении) есть редуکتивная связь. Если мы не признаём существование высокоуровневого кипения, то проблема не возникает, но если мы считаем, что объективно существуют оба кипения, то возникает вопрос – а *почему* наша вселенная организована таким образом, что в ней есть подобная связь между уровнями (уровни признаются реальными). И этот соединительный мост всегда будет ускользать от физикалиста, потому что сам он и должен быть объявлен конечным редуکتивным объяснением. Выход для физикалиста – принятие мереологического нигилизма. Необходимость объяснения межуровневых связей отпадёт в силу декларации их отсутствия.

3) Дилемма Столджара

В работе 2010-го года Дэниел Столджар сформулировал дилемму, с которой неизбежно столкнётся любая версия физикализма. Если физикализм – это утверждение о существовании только физических фактов, то оно ложно так как существуют и другие факты; если же физикализм предполагает что-то более сложное, то становится непонятно, почему это – именно физикализм [Stoljar, 2010, p. 90].

Если редукция – это всегда элиминация, то ложность первой ветви дилеммы очевидна. Конечно, физикалист может настаивать на той или иной форме либерализованной версии своей метафизики [Stoljar, 2010, p. 70]. Однако в последнем случае эта позиция должна быть прочитана как скептицизм – утверждение о том, что мы заявляем, что мир не какой-то, а какой угодно с неясными границами. То что для названия компонентов мира здесь выбрано слово «физическое» – случайность, мы всё равно вынуждены будем заявлять, что каждая новая версия науки должна будет определить, что включается в эту категорию. Так как наука всегда находится в ситуации, когда её положения могут быть персмотрены перед лицом новых свидетельств [Nagel, 2012, p. 3], то подобная версия физикализма достаточно бессмысленна и в познавательном смысле является тождественной самому радикальному скептицизму.

4) Высокоуровневые науки уже автономны

Когнитивно-поведенческая психология успешно моделирует личность без полного обращения к биологии. Социология описывает общество и происходящие в нём процессы без ссылки на психологию. Историки строят гипотезы и проверяют их, не полагаясь на химические процессы, которые несомненно происходили в прошлом. Эти науки эффективно предсказывают и объясняют явления на своём уровне, что даёт доказательства в пользу автономии этих наук. Более того, для выявления законов биологии, совершенно не обязательно знание процессов на химическом уровне. Безусловно, какую-то связь между уровнями науки устанавливают. Однако, по всей видимости, междуровневое взаимодействие может носить

характер не только зависимости, но и причинного взаимодействия, реализации, ограничений, конвергенции процессов и так далее. Законы выявленные частными науками самодостаточны для описания процессов на их уровне.

Эмерджентизм оказывается более правдоподобным, потому что он признаёт автономию высокоуровневых наук и законы, которые действуют на этих уровнях, в то время как редуктивизм не может должным образом учесть эти факты. Эмерджентистская метафизика – более консервативна, эмпирически более эффективна в объяснении реальности, чем жёсткий редуктивный физикализм. Кроме того, она не имеет явных философских проблем, в силу того, что объекты реальности, которые мы в состоянии индивидуировать как таковые, продолжают существовать и обладать силами, которые мы реально наблюдаем.

Литература

Corradini, A., O'Connor, T. (2010). *Emergence in Science and Philosophy*. New York. Routledge.

Gillet, C. (2025). *Reduction, Emergence, and the Metaphysics in Science. Exploring New Foundations*. Cambridge. Cambridge University Press

Nagel, E. (1979). *The Structure of Science. Problems in the Logic of Scientific Explanation*. Cambridge. Hackett Publishing Company.

Nagel, T. (2012). *Mind and Cosmos. Why the Materialist Neo-Darwinian Conception of Nature Is Almost Certainly False*. New York. Oxford University Press.

Stoljar, D. (2010). *Physicalism*. New York: Routledge.

Stoljar, D. (2024). Physicalism. *The Stanford Encyclopedia of Philosophy* [Электронный ресурс] URL: <https://plato.stanford.edu/archives/spr2024/entries/physicalism/> (дата обращения: 27.05.2025).

Varzi, A. (2019) Mereology. *The Stanford Encyclopedia of Philosophy* [Электронный ресурс] URL: <https://plato.stanford.edu/archives/spr2019/entries/mereology/> (дата обращения: 27.05.2025).

К. Е. Морозов

Институт философии РАН, Москва

lovecraft.wittgenstein@gmail.com

Феноменальное сознание – это способность переживать субъективный опыт. Качественные аспекты этого опыта, выражающие то, каково его переживать, называют «квалиа». Часто предполагается, что феноменальное сознание и квалиа создают серьёзные проблемы для материализма, потому что их сложно (если вообще возможно) редуцировать к чему-то физическому [Чалмерс, 2013]. Материалисты по-разному отреагировали на этот вызов. Некоторые из них согласились, что феноменальное сознание не удастся редуцировать к каким-то знакомым нам физическим свойствам, но при этом настаивали на необходимой метафизической зависимости сознания от физических процессов [Сёрл, 2015; Stoljar, 2006]. Другие, напротив, настаивали на возможности редуктивного объяснения феноменального сознания [Kim, 2005; Prinz, 2012].

Но некоторые материалисты, которых называют «иллюзионистами» или иногда «элиминативистами», пошли совершенно иным путём. Они утверждают, что феноменального сознания не существует, но у нас есть лишь иллюзия подобного сознания [Frankish, 2017]. Некоторые критики и даже сторонники иллюзионизма подмечали тот факт, что эта позиция сталкивается с некоторыми трудностями в том, что касается объяснения и

обоснования наших моральных убеждений [Strawson, 2019; Kammerer, 2020; Balog, 2023]. Основную проблему эти авторы видят в том, что иллюзионизм не может объяснить неправомерность причинения боли чувствующим существам. Также иллюзионизм не может предложить ясный критерий морального статуса, поскольку наиболее правдоподобным критерием является само феноменальное сознание [Seager, 2001; Sebo, 2017; Lee, 2022].

Акцент на этой проблеме морального статуса, однако, затемняет другую, не менее значимую нормативную проблему иллюзионизма. Дело в том, что феноменальное сознание выполняет ещё одну особую нормативную роль – обеспечивает нас основаниями для действий, т.е. участвует в моральной мотивации. Моральная мотивация – это частный случай нормативной мотивации, который выражается в том, что, как правило, моральные суждения имеют для нас мотивирующую силу. Если я утверждаю, что морально неправильно мучить кошек, то это мотивирует меня не участвовать в действиях, предполагающих мучение кошек, а также, возможно, активно противодействовать всем, кто участвует в подобных действиях. Утверждение «Морально неправильно мучить кошек», как кажется, является описательным: оно приписывает действию «Причинение мучений кошкам» свойство «Моральная неправильность». Но другие описательные суждения, вроде «На улице идёт дождь», как правило, не порождают мотивацию к действиям.

Этот особый статус моральных суждений ставит нас перед проблемой объяснения феномена моральной мотивации. Как я утверждаю, одно из правдоподобных объяснений этого феномена

состоит в том, что в моральной мотивации участвует особое качественное специфичное переживание – чувство правильности. Ментально здоровый сознательный агент, когда он сталкивается с ситуациями явной несправедливости, переживает особое чувство, которое мотивирует его поступить в соответствии с требованиями морали. При этом мы не можем истолковать это чувство чисто функционально (как того захотели бы иллюзионисты), потому что за вычетом их феноменальной специфики мы не сможем отличить моральную мотивацию от чисто руденциальной или иррациональной.

Это чувство правильности можно объяснить двумя разными способами, каждый из которых определённым образом согласуется с нашими моральными убеждениями. С одной стороны, мы можем рассматривать это чувство через призму репрезентационалистских теорий сознания [Иванов, 2020; Гусев, 2021]. С этой точки зрения, существуют моральные (шире – нормативные) факты, которые сами по себе обладают определённой мотивирующей силой. И поскольку, согласно репрезентационализму, феноменальное сознание репрезентирует свойства внешнего мира, чувство правильности является просто формой ментальной репрезентации этих нормативных фактов. Такое понимание чувства правильности близко кантовской идее «благоговения перед моральным законом».

Альтернативный подход связывает мотивирующую силу моральных суждений с нашей способностью к эмпатии [Zahavi, 2014; Marshall, 2018]. С этой точки зрения, мораль сущностно связана с нашей способностью проецировать на себя субъективную перспективу другого агента. Поскольку собственные страдания в силу своей феноменальной специфики

мотивируют нас предпринимать нужные действия, чтобы избежать или остановить эти страдания, а эмпатия позволяет нам разделить субъективную перспективу другого агента, эмпатические переживания также имеют для нас мотивирующую силу. Эти два объяснения также необязательно рассматривать как взаимоисключающие.

Чувство правильности является одним из ключевых аспектов нашей феноменологии моральной агентности. Но специфика феноменологии агентности (не только моральной) также состоит в том, что мы не можем полноценно отделить переживание агентности от реализации действий (ср.: [Яшин, 2024]). Именно эта предполагаемая делимость в случае перцептивных и аффективных переживаний, как я полагаю, подпитывает интуиции в пользу иллюзионистской элиминации феноменального сознания. Однако анализ моральной мотивации показывает, что мы не можем столь надёжно отделить феноменальные и функциональные аспекты переживаний, а также не можем объяснить специфику моральной мотивации, вообще не обращаясь к качественным аспектам нашего опыта.

Литература

Гусев, А. А. (2021). Квалиа-реализм и репрезентационализм. *Философия науки и техники*. Т. 26. № 1. С. 107–121. DOI: 10.21146/2413-9084-2021-26-1-107-121.

Иванов, Д. В. (2020). *Природа феноменального сознания*. М. Либроком.

Сёрл, Дж. (2015). *Как можно решить проблему «сознание – тело»?* М. Идея-Пресс.

Чалмерс, Д. (2013). *Сознающий ум: В поисках фундаментальной теории*. М. Либроком.

Яшин, А. С. (2024). Достаточно ли гладко выбрит зомби для работы в метафизике? *Финиковый компот*. № 19. С. 64–68. DOI: 10.24412/2587-9308-2024-19-4-76.

Balog, K. (2023). The Abolition of Phenomena: A Voyage among the Zombies. *Klesis*. No. 55. URL: <https://www.revue-klesis.org/pdf/klesis-55-07-katalin-balog-the-abolition-of-phenomena-a-voyage-among-the-zombies.pdf> (дата обращения: 28.02.2025).

Frankish, K. (ed.) (2017). *Illusionism as a Theory of Consciousness*. Exeter. Imprint Academic.

Kammerer, F. (2020). The Normative Challenge for Illusionist Views of Consciousness. *Ergo: An Open Access Journal of Philosophy*. Vol. 6. No. 32. P. 891–924. DOI: 10.3998/ergo.12405314.0006.032.

Kim, J. (2005). *Physicalism, or Something Near Enough*. Princeton. Princeton University Press.

Lee, A. Y. (2022). Speciesism and Sentientism. *Journal of Consciousness Studies*. Vol. 29. No. 3–4. P. 205–228. DOI: 10.53765/20512201.29.3.205

Marshall, C. (2018). *Compassionate Moral Realism*. Oxford. Oxford University Press.

Prinz, J. J. (2012). *The Conscious Brain*. Oxford. Oxford University Press.

Seager, W. E. (2001). Consciousness, Value and Functionalism. *Psyche: An Interdisciplinary Journal of Research on Consciousness*. Vol. 7. No. 20. P. 1–19.

Sebo, J. (2017). Agency and Moral Status. *Journal of Moral Philosophy*. Vol. 14. No. 1. Pp. 1–22. DOI: 10.1163/17455243-46810046.

Stoljar, D. (2006). *Ignorance and Imagination: The Epistemic Origin of the Problem of Consciousness*. Oxford. Oxford University Press.

Strawson, G.J. (2019). A Hundred Years of Consciousness: “A Long Training in Absurdity”. *Estudios de Filosofía*. Vol. 59. P. 9–43. DOI: 10.17533/udea.ef.n59a02.

Zahavi, D. (2014). Empathy and Other-Directed Intentionality. *Topoi*. Vol. 33. No. 1. P. 129–142. DOI: 10.1007/s11245-013-9197-4.

Мистерианство и аргумент многоуровневой абстракции

А. С. Павлов

Институт философии РАН, Москва

pavlov.alexey@gmail.com

Исследование выполнено за счёт гранта Российского научного фонда (проект No 24–28–00804), <https://rscf.ru/project/24-28-00804/>

Мистерианство (от англ. mystery – «тайна») – это подход к решению проблемы «сознание–тело», предложенный американским философом Колином Макгинном [Павлов, 2023]. Методологические основания мистерианства разработаны Макгинном в контексте *трансцендентального натурализма*. Эту концепцию можно охарактеризовать как версию натурализма, дополненную *тезисом когнитивной ограниченности*: сознание – это естественный феномен, чьи характеристики предопределены эволюцией; отсюда следует допустить наличие у человека – по аналогии с другими видами – естественных когнитивных ограничений. Позиция Макгинна примечательна оппозиционностью по отношению как к дуалистическим, так и материалистическим теориям. С дуалистами Макгинна объединяет признание реальности феноменального сознания, однако он считает, что дуалисты ошибаются, когда говорят о метафизической независимости сознания от мозга и его принадлежности некоей нематериальной субстанции (душе). В

свою очередь, материалистические теории не устраивают Макгинна потому, что, по его мнению, постулируют редуцируемость ментального к физическому, не обеспечивая интуитивным пониманием этих отношений. В целом макгинновская постановка проблемы «сознание–тело» покоится на трёх тезисах: *тезис натурализма* (всё существующее имеет физическую природу), *тезис феноменального реализма* (феноменальный опыт реален) и *тезис метафизической зависимости ментального от физического* (феноменальный опыт порождается мозгом). В центре аргументации Макгинна стоит *тезис когнитивной замкнутости*: человеческий интеллект в том виде, в котором он существует сейчас, не способен решить проблему «сознание–тело». Дополнительно вводится *гипотеза C•–свойства*: существует некий тип когнитивно недоступных для нашего ума свойств, выступающих в качестве онтологических и каузальных посредников между физическими и ментальными свойствами; имея представление о C•–свойствах, мы могли бы иметь интуитивное понимание связи сознания и мозга. Таким образом, *конструктивное решение проблемы «сознание–тело»* фактически представляет собой вариант редуктивного физикализма, обеспечивающего нас интуитивным схватыванием отношений ментального и физического.

Российский ML–разработчик Данила Кутный [Kutniy, 2021] предложил новый аргумент в поддержку мистерианства (*от многоуровневой абстракции*).

Представим нейрон, получающий сигналы от множества связей. Понимание его функционала предполагает расшифровку того, как получаемые данные на входе преобразуются в данные на выходе. Важно учитывать, что выходные данные могут не

отсылать к конкретным паттернам или визуальным компонентам. Вместо этого они могут репрезентировать абстрактный аспект визуальной информации (очертания, цвета, текстуры и т.п.). Таким образом, уже исследование функционирования одного нейрона представляет серьезные трудности. Если допустить, что каждый нейрон действует в пространстве тысяч взаимосвязей, то полное функциональное объяснение нейронных процессов мозга может оказаться непосильной для человеческого ума задачей.

Аналогичным образом дело обстоит с аппаратной частью современного искусственного интеллекта (далее – ИИ). Транзисторы – это базовые компоненты, контролирующие электрические сигналы. Они образуют *логические вентили* (logical gates), выполняющие простейшие логические операции (И, ИЛИ, НЕ и др.) и объединяющие входные данные для производства выходных данных. Над логическими вентилями находятся арифметико-логические устройства, которые интегрированы в центральный процессор (далее – ЦП), организующий поток данных и выполняющий инструкции. Затем следует язык ассемблера, обеспечивающий понятное для человека представление машинного кода и позволяющий программистам писать инструкции для ЦП. На более высоком уровне расположены языки программирования (например, Python или Java), позволяющие создавать сложные приложения с синтаксисом, приближенным к человеческому.

Подобно микропроцессорам, искусственные нейронные сети (далее – ИНС) также работают при помощи многоуровневой абстракции. На базовом уровне находятся *синаптические веса*, представляющие силу связей между нейронами. Каждый нейрон

получает входные сигналы по множеству соединений, и веса определяют, насколько сильно каждый входной сигнал влияет на выходной сигнал нейрона. Над весами находятся сами нейроны, выполняющие нелинейные преобразование входных сигналов. Функция активации каждого нейрона определяет, должен ли он сработать на основе суммы своих входных сигналов, что вносит вклад в общую функциональность сети. Нейроны организованы в слои, где каждый слой преобразует данные во всё более абстрактные представления. Затем выходной слой генерирует окончательную классификацию или прогноз на основе обработанной информации. Каждый слой строится на основе предыдущего аналогично тому, как ЦП объединяет различные компоненты для выполнения сложных задач.

Согласно аргументу многоуровневой абстракции, если мы не можем иметь полного представления о функционировании нейронной сети, то мы не способны объяснить, как она производит сознание. Аргумент поддерживает мистерианство как по отношению к проблеме естественного сознания (далее – ЕС), так и по отношению к проблеме искусственного сознания (далее – ИС).

Кутный оговаривает ограниченность актуальной научной методологии и приводит в качестве примера исследование Джонаса и Кординга [Jonas, Kording, 2017], в котором посредством фМРТ удалось идентифицировать нейронноподобные структуры внутри микропроцессора, соответствующие определённым вычислительным задачам. Тем не менее, исследование не привело к всестороннему пониманию функционирования микропроцессора.

По мнению Кутного, решение проблемы ЕС – более трудная задача. Организация нейронных сетей мозга продиктована эволюцией. Мозг состоит из систем, по своей природе более сложных и трудных для интерпретации, чем микропроцессор. Каждый органический нейрон – это сложная клетка с тысячами синапсов и различными ионными каналами, участвующими в сложных биохимических процессах. В противоположность этому, искусственные нейроны представляют собой упрощённые модели, созданные для выполнения конкретных функций, и часто не способны полностью отразить всю функциональную сложность своих биологических аналогов.

На мой взгляд, аргумент многоуровневой абстракции сталкивается с тремя проблемами. Во-первых, получение полного знания о функционировании нейронной сети может быть нелёгкой задачей, однако неочевидно, что эта задача в принципе не выполнима, и что такого рода знание когнитивно недоступно для человеческого ума. Во-вторых, неочевидно, что полное знание о функционировании нейронной сети позволит решить проблему «сознание–тело»: аргумент явно опирается на редуктивный физикализм, и у противников этой теории найдётся что сказать против. В-третьих, есть философы, сомневающиеся в возможности создания феноменального сознания на неорганической основе (например, Сёрл [Сёрл, 2003, С. 331]), и если они правы, то аналогия между ЕС и ИС работает лишь отчасти. В этом контексте, возможно, имеет смысл постановка вопроса о *поиске признака феноменальности у ИИ* (по аналогии с *поиском признака когнитивности у когнитивных систем* [Павлов, 2024, С. 251–260]).

Литература

Павлов, А. С. (2023). «Новое мистерианство» Колина Макгинна: почему мы не можем решить проблему «сознание–тело?». Москва. URSS.

Павлов, А. С. (2024). Мистерианство на границе философии сознания и когнитивных наук: обзор новых аргументов pro et contra. *Историко-философский ежегодник*. Т. 39. С. 236–265.

Сёрл, Д. (2003). Разум, мозг и программы. Хофштадтер, Д., Деннет, Д. (ред.) *Глаз разума*. Самара. Бахрах-М. С. 314–331.

Jonas, E., Kording, K. P. (2017). Could a neuroscientist understand a microprocessor? *PLoS Computational Biology*. 13(1), e1005268.

Kutniy, D. (2021). The Boundaries of Cognitive Closure: Argument for Mysterianism in the Philosophy of Consciousness [Электронный ресурс]. URL: <https://philarchive.org/archive/KUTTVO> (дата обращения: 27.05.2025)

ФЕНОМЕНАЛЬНЫЕ И ФЕНОМЕНАЛЬНЫЕ СВОЙСТВА

СОСТОЯНИЯ

М. С. Сысоев

Институт философии РАН, Москва

sysoev.paritet@gmail.com

Феноменальное состояние – это такое ментальное состояние, для которого есть то, *каково это* быть в этом состоянии [Block, 1995]. Феноменальные качества – это такие феноменальные свойства, которые характеризуют конкретные аспекты феноменального состояния, ассоциированные с разными модальностями восприятия: цвета, звуки и другие ощущения. Качественные свойства также можно назвать квалитативными свойствами.

Вопрос, который я хотел бы рассмотреть далее, заключается в следующем: как между собой соотносятся феноменальные состояния, т.е. состояния, характеризующиеся тем, *каково это* быть, в данном состоянии, и феноменальные качества, такие как «красный», «боль», «запах яблока» и другие? Чтобы наметить основные ответы на вопрос, поставленный ранее, необходимо определиться с ответом на два подвопроса:

Вопрос №1: Является ли феноменальное состояние *не более чем* экземплификацией какого-то набора феноменальных свойств или же это нечто большее?

Вопрос №2: Являются ли феноменальные качества единственной разновидностью феноменального свойства или же существуют какие-то еще феноменальные свойства?

Касательно вопроса №1 следует отметить, что состояние удобно описывать как экземплификацию свойств даже в том случае, если его природа этим не исчерпывается. Условимся, что состояние может быть идентифицировано *с некоторой степенью точности* путём указания на три параметра (x, P, t): во-первых, объект, который обладает состоянием (x), свойства, которые состояние экземплифицирует (P) и времени, в которое это состояние имеет место (t). При этом поскольку речь идёт о феноменальном состоянии, то необходимо, чтобы (P) было именно феноменальным свойством. Я предполагаю, что с указанным способом идентификации мало кто будет спорить, по крайней мере среди реалистов в отношении феноменального сознания. Однако степень точности такой идентификации будет определяться как раз тем, существует ли что-то еще, значимое для природы феноменального состояния, и здесь уже возможны разные точки зрения. Если такой способ идентификации максимально точен, то ничего больше не требуется. Назовём этот подход *минимализмом*. Если же такой способ идентификации не точен, то требуется что-то еще. Что это может быть? Перечислю некоторые варианты дополнительных ограничений, в скобках помещая проблемный вопрос, который связан с этим вариантом:

- экземплифицируемые феноменальные свойства должны быть представлены в определённом количестве (будет ли феноменальным состояние с одной единственной экземплификацией одного феноменального свойства?)
- среди экземплифицированных свойство должны быть определённые качественные свойства (будет ли феноменальным состояние, в котором многократно

экземплифицировано одно и то же феноменальное свойство?).

- необходимы какие-то особые отношения между свойствами, например, структурные, (будет ли феноменальным состояние, экземплифицирующее любой набор никак не связанных друг с другом свойств?).
- необходимы какие-то специфические свойства, которые не сводятся к качественным, например, свойство субъективности, приватности, единства [Вауне, 2009] и т.д. (является ли феноменальным состояние, которое обладает качествами, но не обладает субъективностью, единством? Существуют ли вообще такие состояния? Или же никаких дополнительных свойств не требуется?).

На первый взгляд перечисленные варианты могут показаться несколько странными, если мы принимаем предпосылку, что любая экземплификация феноменального свойства делает экземплифицирующее его состояние феноменальным. Но принятие этой предпосылки необходимо лишь при принятии минимализма, у нас нет никаких оснований заранее предрешать этот вопрос. Может оказаться, что феноменальные состояния концептуально первичны и феноменальные свойства – это лишь результат процедуры абстракции условных частей таких состояний. В этом случае экземплификация какого-либо отдельно взятого феноменального свойства – это не достаточное условие для существования феноменальных состояний.

Я утверждаю лишь то, что у нас нет хороших оснований для принятия минимализма. Мне не известны способы, позволяющие легко устранить перечисленные выше варианты. Единственное

возражение, которое, как я полагаю, могло бы заблокировать рассмотрение этих вариантов, заключается в следующем тезисе, ограничивающем понятие феноменального:

(ОФ) всё, что не имеет отношения к феноменальным качествам – не касается феноменального сознания, качественность (квалитативность) состояний и свойств – это и есть их феноменальность.

Данный тезис предполагает, что вопросы о структуре феноменальных свойств, отношений между этими свойствами, их количества и т.д. – это вопросы, связанные с функциональным аспектом психики, а не с феноменальным сознанием. Проще говоря – от феноменального аспекта требуется только предоставить набор качеств, а остальное за него выполняют функциональные структуры. В актуальной дискуссии мало кто придерживается подобной точки зрения, поскольку кроме качественности феноменального аспекта всегда принято говорить, как минимум, о субъективности и приватности (Туге, 2021). Для реалиста в отношении феноменального сознания, принимающего ту или иную версию дуализма свойств, данный ход кажется весьма спорным потому, что каузальная релевантность сознания оказывается достаточно слабой. Допустим, что функциональная структура организует качества в зрительном поле определённым образом по каким-то сугубо функциональным причинам, не связанным с итоговой феноменальной картиной. Тогда появление красно-зеленого округлого объекта (яблока) в левой части моего зрительного феноменального поля ничего не выражает и ни на что не влияет, а является лишь феноменальным довеском к тем вычислениям, которые происходят где-то в мозге с участием функциональной

репрезентации яблока и репрезентации моего собственного тела в пространстве. Чтобы сохранить каузальную релевантность феноменального сознания, мы должны признать, что структура качеств имеет не только функциональное, но и феноменальное выражение.

Переходя к вопросу № 2, мы можем обнаружить, что ранее сделанные промежуточные выводы по вопросу №1 непосредственно на него влияют.

Если имеют место альтернативы (3) и (4), то одних только качественных свойств для возникновения феноменального состояния будет недостаточно. Это означает, что существуют иные, не-качественные или не-квалитативные феноменальные свойства, например, свойства приватности, субъективности, единства, а также отношения между свойствами и структуры таких свойств, наличие или отсутствие которых повлияет на идентификацию состояния в качестве феноменального. Резонный вопрос, который здесь может возникнуть, звучит так: почему не включить все эти дополнительные свойства в разряд квалитативных, предполагая, что с ними связано особое качественное переживание. Это интересный вопрос для классификации феноменальных свойств, но на суть дела он не влияет. Если мы принимаем, что существует такой широкий класс феноменальных качеств, то нам придётся выделить особые подклассы: один, отвечающий за репрезентацию конкретных свойств, а второй – отвечающий за структуры свойств, отношения между такими свойствами и т.д. Нам также придётся выделять среди феноменальных качеств такие, которые репрезентируют объекты, и такие, которые репрезентируют отношения между самими репрезентациями. Так или иначе, это

усложняет привычный разговор о феноменальных свойствах. «Красный» уже не может служить парадигмальным примером для всего класса феноменальных свойств, а феноменальное состояние не может являться просто экземплификацией произвольно взятого набора феноменальных качеств. Таким образом, оба возможных ответа на вопрос №2 совместимы с отрицанием минимализма в ответе на вопрос №1.

Литература

Block, N. (1995). On a Confusion About a Function of Consciousness. *Behavioral and Brain Sciences*. Vol. 18(2). P. 227–247.

Bayne, T. (2009). *The Unity of Consciousness*. Oxford University Press.

Tye, M. (2021) Qualia. *The Stanford Encyclopedia of Philosophy* [Электронный ресурс]. URL: <https://plato.stanford.edu/archives/fall2021/entries/qualia/> (дата обращения: 10.05.2025).

Влияние интеллектуальных технологий на общество цифровизации технологий

ПРОЦЕССЫ ЦИФРОВИЗАЦИИ И СМАРТИЗАЦИИ В ДИСКУРСЕ СОВРЕМЕННОЙ ОТЕЧЕСТВЕННОЙ КУЛЬТУРЫ

В. И. Баринов

МОУ «Ряжская СШ №4», Ряжск
sedriksakson@gmail.com

На современном этапе развития общества наблюдается эпоха динамичных трансформаций, обусловленных существенным воздействием техногенной культуры на социальную структуру. Прогрессирующая глобальная цивилизация продвигается вперёд ускоренными темпами вследствие появления и непрерывного внедрения инновационных технологических решений, оказывающих влияние на все ключевые сферы жизнедеятельности человека. В этой связи необходим комплексный анализ структуры современной техногенной культуры, который позволит раскрыть совокупность значимых факторов, оказывающих фундаментальное влияние на протекающие социокультурные трансформации. Особое значение здесь приобретают феномены цифровизации и смартизации, представляющие собой

основополагающие тенденции, обуславливающие радикальные изменения в организации повседневной активности общества.

Цифровизация сегодня представляет собой естественный этап развития современного общества, который органично продолжает два предшествующих этапа преобразования социума: процессы информатизации и компьютеризации. Однако цифровизация выходит далеко за рамки простого продолжения этих явлений и становится самостоятельным этапом.

Следует отметить, что основы современной информационной инфраструктуры и цифровой экономики XXI века закладывались ещё в начале XX века. Процесс информатизации, начавшийся тогда и связанный с появлением электронно-вычислительных машин, сопровождался созданием и распространением информационных ресурсов: информационных баз данных, электронных библиотек и справочных систем. Важнейшей задачей того времени стало развитие доступности этих ресурсов для пользователей во всём мире и формирование у человека технологий управления большими объёмами информации, представленной в виде данных, что впоследствии привело к формированию информационной культуры.

Процесс компьютеризации, запущенный после широкомасштабной кампании информатизации общества, обусловлен прежде всего появлением пятого поколения электронных вычислительных машин, основанных на базе центральных процессоров. Он обозначил радикально новую фазу эволюции информационно-коммуникационных технологий, характеризуемую значительным расширением доступности высокопроизводительной вычислительной техники среди широких слоев населения.

Компьютерная модернизация сопровождалась экспоненциальным увеличением количества персональных компьютеров, которые получили повсеместное распространение – от крупных промышленных предприятий и организаций государственного сектора до индивидуальных домохозяйств. Массовое внедрение персональной вычислительной техники превратилось в неотъемлемый атрибут жизнедеятельности представителей различных профессиональных сообществ и социально-демографических групп, что заложило основу для формирования компьютерной культуры.

Процесс цифровизации, являясь закономерным продолжением двух предыдущих этапов, отличается от них запуском цифровой трансформации, которая охватывает более широкий спектр изменений и носит полиморфный характер. Это означает, что цифровизация одновременно воздействует на различные сферы жизни общества, преобразуя их структуру и механизмы функционирования, что приводит к формированию цифровой культуры.

Цифровая культура в современном информационном пространстве охватывает все сферы жизнедеятельности человека, радикально трансформируя экономику, политику, социальные отношения, способы общения, личные контакты и мировоззрение значительной части населения [Горлова, 2020].

Тем не менее представляется целесообразным рассматривать такие явления, как информатизация, компьютеризация и цифровизация, а также возникшие благодаря этим процессам типы культур – информационную, компьютерную и цифровую, входящие в структуру техногенной

культуры, именно как феномены, сформировавшиеся и развившиеся в условиях третьей техно-революции.

Однако сегодня человечество вступило в принципиально новый этап своего цивилизационного развития – четвёртую техно-революцию, отличающуюся от предыдущей своей масштабностью, скоростью распространения технологий и глубоким воздействием на все сферы жизни общества [Шваб, 2016]. Именно она породила совершенно новые продукты и технологии – смарт-устройства, а также смарт-технологии, запустившие новый этап развития общества, в котором сформировались необходимые условия для становления и развития смарт-культуры. Данный феномен сопровождается процессом, который мы называем «смартизация».

Процесс смартизации подразумевает применение инновационных технологий интеллектуального типа, ориентированных на совершенствование всех аспектов существования и трудовой деятельности человека. К таким смарт-технологиям относятся интернет вещей, большие данные, искусственный интеллект и нейросети, которые в комплексе со смарт-устройствами образуют структуру смарт-дома и смарт-города. Это позволяет существенно повышать продуктивность функционирования муниципальных услуг, транспортных сетей, жилищно-коммунального хозяйства, производственного сектора и сервисных предприятий. Примером служит возможность удаленного мониторинга городской инфраструктуры, контроля уровня ресурсопотребления, оперативного реагирования на кризисные ситуации через интегрированные IoT-решения.

И процесс цифровизации, и процесс смартизации кардинальным образом преобразуют архитектуру существующих

общественных связей, порождают новую культурно-цивилизационную парадигму и задают вектор эволюции коммуникативных форматов среди членов общества. Цифровая среда интенсифицирует сетевое взаимодействие пользователей, создает условия для возникновения цифровой экосистемы, облегчающей обмен опытом и сведениями. Параллельно смартизация существенным образом революционизирует привычные механизмы хозяйствования, форсирует научно-технический прогресс, воздействует на качественные характеристики рабочей силы и профессиональный профиль занятых лиц.

Следовательно, оба процесса выступают ключевыми движущими силами становления новейшей действительности, видоизменяя поведенческие паттерны, бытие и ежедневные практики миллионов человек, непосредственно воздействуя на состояние социально-экономической среды, экономической динамики и политического устройства общества.

Литература

Горлова, И. И. (2020). Цифровая культура в информационном обществе. И. И. Горлова, А. Л. Зорин (ред.). *Культурное наследие России*. № 2 (29). С. 3–9. DOI 10.34685/NI.2020.29.2.001.

Шваб, К. (2016). *Четвертая промышленная революция*. М.

ТРАНСФОРМАЦИЯ СОЦИАЛЬНОЙ КОММУНИКАЦИИ В ХОДЕ РАЗВИТИЯ ТЕХНОЛОГИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Г. В. Вышедко

Институт Философии НАН РБ, Минск

gleb.vv7@gmail.com

*Статья подготовлена при финансовой поддержке БРФФИ,
проект № Г25ИИ М-014 «Технологии искусственного
интеллекта и мировоззренческие структуры: ключевые
аспекты взаимовлияния»*

В данной статье рассматривается феномен трансформации социальной коммуникации в ходе развития технологии искусственного интеллекта. Отмечается возрастание роли влияния систем ИИ как на индивидуальную, так и на социальную коммуникацию. ИИ становится важным агентом взаимодействия между ИИ и человеком, предлагая последнему поддержку, интерактивность, помощь. В социальном разрезе ИИ становится фактором социальных изменений в системе коммуникации, формируя процесс передачи, формирования и аккумуляции информации. В результате из-за неравномерного и несправедливого распределения и распространения информации существует риск возрастания социальной нестабильности. Разумное и постепенное внедрение достижений в области ИИ может помочь смягчить данное обстоятельство.

Современное общество часто характеризуется как турбулентное и неустойчивое. Одними из источников социальных изменений становятся прорывные технологии, временами радикально меняющие характер социальной коммуникации. Одной из таких технологий в XXI веке, радикально поменявшая статус социальной коммуникации стала технология ИИ. Системы ИИ как интеллектуальные системы, обладающие многими чертами человеческого разума такие как общение, рассуждение, решение проблем и т.д. во многом расширили способности человека [Dzyaloshinsky, 2021]. ИИ может в мгновение секунды сделать краткий пересказ большой статьи, сделать обзор книги, провести краткое исследование на определенную тематику – все это расширяет возможности коммуникации между людьми. Более того, сам ИИ становится участником социальных коммуникации, выступая не только как посредник, но и как непосредственный соучастник общения. В частности, 10% молодых американцев считают ИИ своим близким другом [Kokotajlo et al., 2025]. А Российский ВЦИОМ в 2023 году провел исследование, в котором обнаружилось, что 55% респондентов доверяет ИИ [ВЦИОМ, 2022]. Такие показатели не являются случайностью, но свидетельством вхождения ИИ в процессы социальной коммуникации. Приостановка совершенствования CHATGPT лишь показывает, насколько дает мимолетное представление становление эпохальной технологии. В данной статье мы постараемся доказать данный тезис.

Из-за своих обширных возможностей системы ИИ могут применять по-разному. Нельзя не обратить внимание на позитивный потенциал систем ИИ. В 2023 году Д. Вудхауз и А.

Чарльзфорс в сентябре 2023 года на основе данных сделали вывод, что СНАТ-GPT может предугадывать будущие решения по процентной ставке. Они использовали ChatGPT для классификации публичных заявлений членов комитета по монетарной политике Банка Англии на ястребиные и голубиные, после чего проанализировали, помогает ли это дополнительное знание лучше прогнозировать решения по ставке [Woodhouse, Charlesworth, 2023]. Как отмечают Ф. И. Храмцов и М. С Шибут: «В условиях глобальной цифровизации и перехода к экономике знаний технологии искусственного интеллекта становятся не просто инструментом преобразований, но движущей силой и стратегическим ресурсом конкурентного потенциала национального государства» [Храмцова, Шибут, 2025].

ИИ также изменяет личную коммуникацию. Надо отметить, что ещё в 2017 году компания Google имплементировала функцию «умного ответа», где ИИ-алгоритм советовал, как отвечать на письмо, предлагал свои шаблоны для ответа [Danso, 2023]. С появлением таких чат-ботов как ChatGPT открылась новые возможности личного взаимодействия между чат-ботом и человеком. Чат-бот может отыгрывать множество социальных ролей: учителя, тьютора, друга, психолога, личного врача и иные роли. И нет резонов считать, что количество людей, обращающихся с различными запросами к тем же чат-ботом, которые как им кажется может удовлетворить их потребности уменьшится. Упомянутый выше факт, что 10% молодых американцев считают ИИ своим близким другом, лишь отражает усиливавшуюся тенденцию виртуализации и атомизации общения [Kokotajlo et al., 2025].

Немаловажным изменением стал анализ поведения пользователей сети Интернет при помощи ИИ. ИИ-алгоритмы в интернете анализируют действия пользователей в сети Интернет и предлагают на их основе персонализированный контент, рекомендации, рекламу. В 2020 году компания Tiktok сделала ставку на поведенческие сигналы в деле алгоритмизации контента: досмотры, повторные просмотры, сохранения или факт того, что пользователь поделился конкретным сообщением. Благодаря алгоритмам каждый узнает новость с ракурса, который зацепит именно его, следовательно, пребывание пользователя в соцсетях стало более удобным и персонализированным.

Однако, вместе с тем, как и с многими другими массовыми и эпохальными изобретениями, в обществе вместе с оптимистическим взглядом на то, как изменит ИИ коммуникацию, появляются прямо противоположными вышеназванным тезисам убеждения. Так, есть опасения что будет ИИ будет подвержен цензуре, а сам он станет источником дипфейков (deepfake), ложных фото и видеоматериалов, которые невозможно отличить от реальных. К примеру, информационный портал ИА «Панорама» выпустила ложную новость, в которой было написано, что глава DeepSeek якобы сказал, что они создали свой ИИ на основе советских разработок [Мицкелюнас, 2025]. Несмотря на неправдивость данной информации эту «новость» сначала подхватили региональные СМИ, а следом и по ТВ. ИИ в таком ракурсе предстает не как источник объективной информации, а как агент, подверженный разного рода когнитивным искажениям, предубеждениям. ИИ также оказывается инструментом, используемым в корыстных целях.

ИИ может не просто давать прогнозы социальных изменений, но и с их помощью изменять ткань социальной реальности. ИИ-алгоритмы, рекомендации критикуются за их несправедливую позицию разделения людей на категории, ИИ-алгоритмы критикуют и за то, что с их помощью возможно будет контроль над интернет-пользователями. Таким образом создается информационный кокон. Например, ценитель электромобилей узнает о том, что Дональд Трамп отменил все субсидии для этой сферы, новость будет подана как трагедия с нотками алармизма. А вот ценитель обычных машин может узнать ту же новость, но с позитивным тоном и приложенным видео, где условный маслкар (muscle car) максимально шумно крутится, создавая облако из выхлопов. В таком взгляде ИИ становится источником социальной несправедливости и дифференциации, выдавая различную информацию разным группам пользователей. К тому же, ИИ хоть умеет отыгрывать заданные роли, быть тьютор и учителем и вовсе быть открытым к общению, тем не менее на некоторых интернет-сайтах с подобной тематикой виднеется следующая дополнение: «Это ИИ чат-бот, а не человек. Рассматривайте все, что он говорит как вымысел. Не следует полагает на сказанное им как на факт или совет».

Таким образом возникновение и развитие технологии ИИ имело разностороннее влияние на коммуникацию в социальной сфере: ИИ становится как источником позитивных изменений: удобства пользования, персонализации контента, товарищества, но также и быть причиной негативных тенденции: способствовать генерации ложной и негативной информации, предлагать несправедливую социальную дифференциацию в ходе генерации и предложения персонализированного контента.

Причиной данных обстоятельств стала слишком быстрая и неравномерная реализация достижений в области ИИ в социальную сферу. Мы предлагаем постепенное и обдуманное введение систем ИИ в разные сферы социальной жизни, что позволит легче предсказать многочисленные казусы и негативные ситуации.

Литература

Dzyaloshinsky, I. (2021). Communication Services in the Age of Digital Civilization: Convergence with Artificial Intelligence. *Communication. Media. Design*. Vol. 6, No 3. P. 80–95. URL: <https://cmd-journal.hse.ru/article/view/13409/13223> (дата обращения 31.05.2025).

ВЦИОМ. (2022). Искусственный интеллект: угроза или светлое будущее? *ВЦИОМ Новости* [Электронный ресурс]. URL: <https://wciom.ru/analytical-reviews/analiticheskiiobzor/iskusstvennyi-intellekt-ugroza-ili-svetloe-budushchee> (дата обращения: 31.05.2025).

Мицкелюнас А. (2025). Глава DeepSeek: «Мы создали наш искусственный интеллект на основе советских разработок». *Издание ИА «Панорама»*. URL: <https://panorama.pub/news/glava-deepseek-my-sozdali-nas> (дата обращения: 31.05.2025).

Храмцова, Ф. И., Шибут, М. С. (2025). Технологии искусственного интеллекта в сфере государственного управления Республики Беларусь. *International Journal of Humanities and Natural Sciences*. Vol. 1-4 (100). С. 127–137.

Kokotajlo, D., Alexander, S., Larsen T., Lifland E., Dean R. (2025). *AI 2027*. [Электронный ресурс]. URL: <https://ai-2027.com/> (дата обращения: 31.05.2025).

Danso S., Annan., M, Ntem., M., Kwaku, B., Awudi, B. (2023). Artificial intelligence and human communication: A systematic literature review. *World Journal of Advanced Research and Reviews*, No. 19(01) P. 1391–1403. DOI: 10.30574/wjarr.2023.19.1.1495

Woodhouse D., Charlesworth, A. (2023). Can ChatGPT Predict Future Interest Rate Decisions? *SSRN*. DOI: 10.2139/ssrn.4572831

Концепция четвертой промышленной революции: основные положения и теоретические принципы

А. Ю. Косенков

ГНУ «Институт философии НАН Беларуси»

namenlos94@yandex.by

Статья подготовлена при поддержке Белорусского республиканского фонда фундаментальных исследований, проект №Г24МП-020 «Социально-философские и эпистемологические аспекты развития цифровых технологий: способы классификации и последствия внедрения»

О четвертой промышленной революции сегодня говорится и пишется немало. Новая реальность, формируемая в результате развития современных технологий, воодушевляет и тревожит. И потому удивительно, что работы автора концепции четвертой промышленной революции Клауса Шваба не становятся предметом основательного анализа. Возможно, причина в кажущейся тривиальности его утверждений: восприятие современности как эпохального рубежа в развитии цивилизации лишает тезисы о промышленной революции новизны и оригинальности. Возможно, дело в личности самого К. Шваба, который воспринимается не как авторитетный мыслитель, а скорее как представитель мировой элиты и архитектор новой реальности. Однако большого значения это не имеет: концепция швейцарского экономиста и главы Всемирного экономического форума стала не просто очередной работой о технологиях, а основой для программ инновационного развития,

реализуемых во всем мире. Поэтому работы автора (по крайней мере, две ключевые книги [Шваб, 2018а; Шваб, 2018б]) должны быть предметом рефлексии. При этом особое исследовательское внимание следует уделять не столько ее содержанию – о возможных перспективах развития общества в ходе дальнейшего внедрения актуальных мегатрендов написано достаточно – сколько теоретическим и мировоззренческим принципам, определяющим структуру, основные идеи и место концепции в интеллектуальном пространстве.

В своих работах К. Шваб утверждает о том, что общество находится на пороге масштабной трансформации, обусловленной развитием интеллектуальных систем, беспилотного транспорта, робототехники, интернета вещей, биоинженерии и прочих технологий. Результатом революции, если описывать идеи автора кратко, станет формирование новых структурно-организационных моделей, стимулирование спроса в ходе включения посредством технологий новых потребителей, трансформация структуры занятости, обусловленная, в первую очередь, автоматизацией труда. Промышленная революция кардинально изменит политическую сферу общества, способствуя появлению альтернативных центров принятия решений, изменению природы вооруженных конфликтов из-за стирания границ между состояниями войны и мира и противостоянием в виртуальном пространстве. Новые технологии в перспективе способны изменить человека, его идентичность, характер отношений с другими субъектами.

Вместе с тем масштабные трансформации сопряжены с рисками: неравенством, экологическими проблемами, угрозой приватности индивида и пр. Путь их преодоления –

реформирование институциональных систем, а также формирование «нового мышления». Правда, в реформаторских разделах работ (а один из них носит весьма нескромное название – «На пути к культурному ренессансу») автор ограничивается лишь общими призывами к коллективной ответственности, взаимному доверию, сотрудничеству, рассуждениями о необходимости новых образовательных программ и инвестиционных вложениях [Шваб, 2018а, с. 132–137; Шваб, 2018б, с. 56].

Взвешенные размышления о негативных и позитивных последствиях промышленной революции позволяют определить К. Шваба как технореалиста (и не согласиться с авторами, которые считают его оптимистом [Ларина, 2018, с. 29]). Автор полагает, что технологическое развитие – необратимый процесс, приводящий к амбивалентным последствиям; ввиду этого необходимо использовать позитивные возможности, предоставляемые технологиями, параллельно преодолевая сопутствующие вызовы. Показательно, что размышляя в одном из интервью о прозрачности, формируемой в результате цифровизации, экономист избегает категоричных оценок, предлагая научиться жить в новых реалиях [L'interview de Klaus Schwab].

Реформаторские заявления К. Шваба демонстрируют его приверженность принципу взаимодетерминации технологий и общества: технологии формируют социум (и способствуют его переходу на новую стадию развития), а развитие технологий определяется социальными факторами. Так, рассуждая о необходимости реформ, экономист обращает внимание на то, что технологии не являются автономными сущностями: «мир

технического» вплетен в «мир социального», ввиду чего он управляет [Шваб, 2018б, с. 43-49]. Технологии социальны по своей природе, являясь не нейтральными инструментами, а отражением установок, ожиданий и целей социальных субъектов. Правда, К. Шваб не подкрепляет тезисы аргументами: ему важно, судя по всему, не столько убедить читателя в истинности тезисов, сколько в их соответствии реалиям постреволюционного общества. Ведь если считать технологии автономными, следует из рассуждений автора, их невозможно контролировать; а если их рассматривать как нейтральные инструменты, то трудно привлекать к ответственности разработчиков и бизнесменов.

Также заметим, что К. Шваб балансирует в повествовании между двумя временными измерениями (что в целом характерно для исследователей цифровизации). Говоря о настоящем технологий, он также смотрит в будущее, что размывает границы между модальными измерениями, побуждает считать предвосхищаемое как актуальную данность. Собственно, о революции он пишет скорее как о процессе, который набирает обороты сегодня, но действительно масштабным и заметным станет в ближайшем будущем. По сути, идеям автора приговор может вынести не сторонний критик, а только само время. Вместе с тем и без центрального тезиса о начале революции, что примечательно, книга не теряет ценности – сложно спорить с тем, что искусственный интеллект, виртуальная реальность и прочие технологические тренды изменят жизнь человека и общества.

В завершение заметим, что работы К. Шваба являются содержательными и актуальными. Экономист в своих работах не уходит в предельно абстрактные рассуждения о революции: он говорит об известных технологиях и рисках их внедрения,

включая в повествование мнения экспертов. Но вместе с тем концепция носит скорее обобщающий характер – автор в работах отражает повсеместно обсуждаемые проблемы, что ставит вопрос об их оригинальности, отводит ему лишь роль ретранслятора, посредника между специалистами и читателями. Также работы автора слабо включены в систему социально-гуманитарного знания: его основные источники – экономическая литература и отчеты международных организаций, прежде всего, ВЭФ. Показательно, что даже рассуждая о ценностях, он ссылается не на работы в области этики, а на отчет возглавляемой им организации [Шваб, 2018б, с. 50–51]. В некоторых разделах книг автор отождествляет понятия общества, политики и системы, что приводит к концептуальной путанице. Понятия революции и промышленной революции им и вовсе не проясняются. Наконец, и концептообразующая идея экономиста о промышленной революции может быть подвергнута критике. Конечно, он пишет, что по мнению других теоретиков сегодняшние трансформации являются продолжением предыдущих революций [Шваб, 2018а, с. 10–11]. Однако К. Шваб не очень убедительно отвечает оппонентам, обращая внимание на высокую скорость трансформаций, их потенциальную широту и глубину как маркеры приближающейся революции.

Из вышесказанного можно заключить, что четвертая промышленная революция – это не только концепция, отражающая восприятие современности как эпохального рубежа в развитии общества. В работах К. Шваба формулируются образы настоящего и будущего, отражаются теоретические принципы и позиции экономиста, которые, ввиду его влияния на

элиты, могут стать доминирующими в мышлении и деятельности субъектов, определяющих ключевые векторы инновационного развития. В таком случае критическое осмысление концепции может способствовать ее достраиванию и совершенствованию. Однако рефлексия над идеями К. Шваба может иметь и другую цель – поиск альтернатив проекту четвертой промышленной революции, формулирование более совершенных стратегий социально-технологического развития. Правда, эти поиски следует начать с признания того, что концепция экономиста далека от идеала, а промышленная революция – не данность, а только один из сценариев развития современного общества.

Литература

Шваб, К. (2018). *Четвертая промышленная революция*. Москва. 208 с.

Шваб К., Дэвис Н. (2018). *Технологии Четвертой промышленной революции*. М. 320 с.

Ларина Е., Овчинский В. (2018). Парадокс Ферми и угрозы будущего. *Изборский клуб*. № 7. С. 28–45.

L'interview de Klaus Schwab. URL: <https://www.youtube.com/watch?v=IJcey1PPiIM&list=WL&index=2&t=7s> (date of access: 19.05.2025).

ИНТЕГРАЦИЯ VR/AR В ТРАДИЦИОННОЕ ОБРАЗОВАНИЕ: РАЗВИТИЕ КОМПЛЕКСНЫХ КОМПЕТЕНЦИЙ И АДАПТИВНЫХ СТРАТЕГИЙ

А. В. Кузнецова

Волжский политехнический техникум, г. Волжский
annavitakuznetsova@yandex.ru

Цифровая трансформация общества предъявляет новые требования к образовательным практикам, включая массовое внедрение технологий виртуальной (VR) и дополненной реальности (AR) в учебный процесс. С одной стороны, VR/AR-решения открывают возможности для глубокой вовлеченности и персонализации обучения [Koumprouros, 2024], с другой – важно сохранить личностно-ориентированные, рефлексивные и социальные компоненты традиционного образования. В данной статье представлена разработанная автором методика интеграции VR/AR с групповыми формами рефлексии и проектным обучением, направленная на формирование «гибких» навыков (soft skills) и комплексных компетенций учащихся. Отмечены ожидаемые преимущества такого подхода: повышение мотивации и креативности учеников, развитие критического мышления и коммуникативных умений, а также адаптивность образовательного процесса через персонализированные задания. Кратко описан планируемый пилотный проект по апробации методики в условиях организации дополнительного образования детей.

Современная эпоха характеризуется быстрым развитием информационных технологий: виртуальная (VR), дополненная (AR) и смешанная реальность (MR) получили широкую популярность как в профессиональной подготовке, так и в образовании [там же]. Новое поколение обучающихся – «digital natives» – привыкло к технологически насыщенным средам, что повышает интерес к использованию VR/AR в школе и вузе. Вместе с тем базовые принципы педагогики, подчёркивающие личностно-ориентированное обучение, рефлексию и социальное взаимодействие, остаются актуальными. Поэтому актуальна задача интеграции VR/AR в традиционный учебный процесс с сохранением личностных и социально-коммуникативных компонентов.

Исследования подтверждают, что VR/AR-технологии способны существенно обогащать обучение и развивать у учащихся «гибкие» и цифровые компетенции. VR-среда обеспечивает высокий уровень погружения и интерактивности, стимулируя концентрацию внимания на предмете [Там же]. Интеграция VR в учебную практику повышает вовлеченность и мотивацию учеников, улучшает понимание материала и способствует развитию творческого и критического мышления. Páidicé Horváth отмечает, что персонализированные 3D-VR среды обучения позволяют значительно повысить эффективность усвоения материала: в её эксперименте студенты, обучавшиеся в VR, показали на 20% лучшие результаты тестов по сравнению с традиционными методами [Horváth, 2021].

Технологии AR дополняют эти эффекты, предлагая обогащённый контекст реального мира. Koumpourgos подчёркивает множество образовательных преимуществ AR:

высокую вовлеченность учащихся, возможность визуализировать сложные абстрактные концепции, персонализацию контента. Так, с помощью AR можно превратить уроки в интерактивную игру, стимулируя критическое мышление и сотрудничество, а также обеспечить индивидуальную подстройку материала под потребности каждого ученика [Koumprouros, 2024].

Современные подходы к образованию требуют развития у учащихся не только предметных, но и междисциплинарных навыков (компетенций 21 века) – коммуникации, командной работы, гибкости мышления, креативности. При этом VR/AR-технологии, вопреки опасениям об изоляции, могут поддерживать совместное обучение и социальное взаимодействие. Например, VR эффективно способствует коллективному обучению: оно мотивирует учащихся, создаёт междисциплинарные пространства для совместной деятельности и развивает социальные навыки [van der Meer N., 2023]. Таким образом, интеграция VR/AR открывает новые возможности для развития комплексных компетенций при условии грамотной педагогической поддержки.

Предлагаемая автором методика сочетает преимущества традиционных и цифровых подходов: VR/AR-технологии выступают дополнением, а не заменой базовой учебной практики. Основные компоненты методики включают следующее:

Во-первых, опытное погружение. Учебные модули с применением VR/AR разрабатываются по темам (например, виртуальные лаборатории, экскурсии, симуляции

профессиональных ситуаций). Такие модули стимулируют активное изучение и понимание материала [Horváth, 2021].

Во-вторых, групповая рефлексия. После VR/AR-сеанса учащиеся объединяются в малые группы для обсуждения опыта. Рефлексия позволяет осмыслить пережитое, выявить затруднения, сравнить виртуальный опыт с реальной практикой. Групповая дискуссия и совместный анализ содействуют развитию критического мышления, самооценки и эмпатии [Журавлев, 2011].

В-третьих, проектное обучение. Интегрированные проекты связывают виртуальные задания с реальными практическими задачами. Например, на занятиях ученикам предлагается создать совместный командный проект (простое приложение или реальную модель). Такой подход способствует развитию навыков планирования, ответственности и командного взаимодействия.

В-четвертых, Soft skills и адаптивность. Во время работы над VR/AR-проектами ученики тренируют коммуникативные навыки, креативность и гибкость мышления. Педагог как фасилитатор акцентирует внимание на взаимодействии внутри группы и взаимном обучении. Поскольку VR-среда позволяет адаптировать задания под разные стили обучения, методика обеспечивает адаптивный характер процесса: сложность виртуальных сценариев постепенно повышается с учётом уровня учащихся.

Таким образом, авторская методика стремится объединить технологии и личностно-ориентированное обучение. VR/AR используются для глубокой практической визуализации и мотивации, а групповая рефлексия и проектная деятельность – для сохранения социальных и рефлексивных составляющих

обучения. По мнению автора, именно такое сочетание позволит формировать устойчивые комплексные компетенции и готовить учащихся к быстро меняющимся условиям (адаптивным образовательным стратегиям).

В рамках валидации методики планируется пилотный проект в детском технопарке «Кванториум «Волжский Политех»» на программе обучение «Создание приложений виртуальной и дополненной реальности»: несколько групп будут работать по разработанной программе с включением VR/AR. Ожидается, что такая организация обучения продемонстрирует рост заинтересованности учащихся, улучшение результатов по отдельным предметам и более активное взаимодействие в группе. Предварительные наблюдения указывают на позитивную динамику развития навыков сотрудничества и самостоятельного решения задач (данные пилотного этапа будут опубликованы отдельно).

Цифровизация образования открывает новые возможности, но предъявляет вызовы для традиционной педагогики. Интеграция VR/AR в классную комнату должна быть сбалансирована с сохранением личностно-ориентированного и социального подходов к обучению. Представленная методика показывает, как сочетать технологическую новизну и проверенные практики: виртуальный опыт обогащается групповыми обсуждениями и проектной работой. С помощью VR/AR существенно расширяются границы адаптивности и междисциплинарности, что способствует развитию критического мышления, креативности и коммуникации у учащихся. В целом, грамотное совмещение традиционного и цифрового обучения может дать мощный синергетический эффект, создавая условия

для формирования целостных, адаптивных образовательных стратегий и комплексных компетенций у нового поколения школьников.

Литература

Журавлев, А. Л., Нестик, Т. А. (2011) Групповая рефлексивность: основные подходы и перспективы исследований. *Знание. Понимание. Умение*. № 3. URL: <https://cyberleninka.ru/article/n/grupповaya-refleksivnost-osnovnyue-podhody-i-perspektivy-issledovaniy> (дата обращения: 25.03.2025).

Карпов, А. В. (2004). Психология рефлексивных механизмов деятельности. М.

Левинтов, А. (2021). Рефлексия. Памяти Владимира Лефевра. М.

Horváth, I. (2021). An Analysis of Personalized Learning Opportunities in 3D VR. *Frontiers in Computer Science*. URL: <https://www.frontiersin.org/journals/computer-science/articles/10.3389/fcomp.2021.673826/full> (дата обращения: 15.04.2025)

Koumpouros, Y. (2024). Revealing the true potential and prospects of augmented reality in education. *Smart Learning Environments*. [Электронный ресурс]. URL: <https://slejournal.springeropen.com/articles/10.1186/s40561-023-00288-0#:~:text=,interest%20in%20the%20subject%20matter> (дата обращения: 21.15.2025).

van der Meer, N., van der Werf, V., Brinkman, W.-P., Specht, M. (2023). Virtual reality and collaborative learning: a systematic literature review. *Frontiers in Virtual Reality*. [Электронный ресурс]. URL: <https://www.frontiersin.org/>

journals/virtual-reality/articles/10.3389/frvir.2023.1159905/full (дата обращения: 21.15.2025).

ПРОГРАММНОЕ ОБЕСПЕЧЕНИЕ ДЛЯ ОПРЕДЕЛЕНИЯ КОЭФФИЦИЕНТА СТРЕССА И ФАЗЫ ПРОФЕССИОНАЛЬНОГО ВЫГОРАНИЯ

С. В. Лихачева, А. А. Попков, А. В. Донцов

Сибирский государственный университет путей сообщения,

Новосибирск

luntik.neko.anime@mail.ru

Киберпсихология – раздел психологии, изучающий взаимодействие человека и цифровой техники, их влияние друг на друга. Данные исследования находят практическое применение как в сфере психологии, так и в сфере информационных технологий [Сайбер, 2010]

Актуальность раздела киберпсихологии, связанного с прогнозированием, в условиях мира информационных технологий постоянно растет. Также возрастает ритм жизни людей, что добавляет факторов, влияющий на психическое состояние человека. По данным исследования Русской Школы Управления за 2024 год, 83% россиян сталкивались с профессиональным выгоранием, а 45% из них находятся в этом состоянии прямо сейчас. Выгорание снижает продуктивность работы сотрудников. [Григорьева, 2024] Эту проблему можно решить с помощью прогноза выгорания сотрудников до выявления этого факта, чтобы принять соответствующие меры по решению данного вопроса. эффективным инструментом для управления здоровьем персонала и повышения общей эффективности работы компании может стать внедрение

программного обеспечения (ПО) для оценки и прогнозирования уровня стресса.

Выгорание – это состояние эмоционального, физического и психического истощения из-за чрезмерного и длительного стресса.

Профессиональное выгорание – это состояние, когда человек чувствует себя крайне уставшим и лишённым энергии и мотивации из-за постоянного стресса на работе.

Отметим, что ранее уже поднимался вопрос об определении уровня выгорания с помощью ПО. Например, такие как Erudit AI и Urlevel, которые оценивают уровень стресса сотрудника, а также его продуктивность на рабочем месте, анализируя частоту его взаимодействия с техническими устройствами, или Aquant's Employee Burnout Prediction Model – модель, которая использует искусственный интеллект для мониторинга ключевых метрик, таких как количество дней с последнего отпуска и ежедневное выполнение задач, чтобы выявлять сотрудников, подверженных риску выгорания.

Таким образом, целью данного исследования является изучение методов прогнозирования выгорания сотрудников с помощью ПО. Для этого необходимо решить ряд задач: определить факторы выгорания, рассмотреть возможность помещения этих факторов в математическую модель для последующего расчета, рассмотреть ПО, которые могут реализовать данный метод.

Факторы, влияющие на выгорание, можно разделить на 2 категории: зависящие от личных качеств сотрудника, а также независимые стресс-факторы.

Начнем с факторов, которые мы считаем зависимыми от сотрудника. Важно отметить, что на этапе составления технического задания и написания программы сотрудник не является человеком, как таковым. Под понятием сотрудника мы имеем ввиду экземпляр класса `Worker`, который хранит в себе некий набор данных, который и будет анализировать наше ПО. Получается, что к данным факторам относятся такие качества, как удовлетворенность работой и коэффициент полезного действия от выполняемой деятельности, так как для позитивной собственной оценки необходимо видеть прогресс, типа нервной системы человека, а также гендерный и возрастной факторы.

Под удовлетворенностью работой мы понимаем то, как человек оценивает свою работу за короткий промежуток времени. По его отчету в свободной форме проводится NLP-анализ (анализ натурального лингвистического процесса). Если рассматривать готовые NLP-модели, то наиболее подходящими можно назвать `Peakon` и `Viva Insights`, которые уже можно встретить в области HR-аналитики, а также `IBM Watson` и `MonkeyLearn`, которые способны сравнивать эмоциональную нагрузку в текстовых отчетах.

Отдельно выделяем дополнительные стресс-факторы. Они не зависят от самого сотрудника, на выбранной должности, а потому мы рассчитываем его в программе, называя коэффициентом стресса, факторы которого: рабочая нагрузка - баланс рабочих часов и часов отдыха, уровень профессиональной нагрузки - объемы работы и наличие дедлайнов, условия труда, коэффициент социальной поддержки и уровень коммуникации на рабочем месте, в целом, технический фактор, возможность карьерного роста, как мотивация к развитию в работе, а также

внешние факторы, такие как заработная плата и стабильность рынка труда.

Ниже представлен способ расчета, основанный на тесте выгорания Бойко [Бойко, 1999].

Данные факторы выводятся в математическую модель линейной вероятности таким образом, что в ней будут суммироваться все факторы перемноженные на их силу влияния в доле от единицы. Среди них силу 0,5 будет иметь фактор коэффициента стресса, который мы объединяем при отдельном расчете, то есть в группе стресс-факторов, не зависящих от сотрудника. Остальные факторы имеют следующие силы: коэффициент удовлетворенности работой – 0,17; тип нервной системы – 0,2; пол и возраст – 0,13.

Таким образом мы получим показатель величиной от -1 до 1, который в последующем будем подставлять в формулу логистической вероятности. Это нам даст показатель, который мы сможем сравнить со шкалой фаз выгорания Бойко, распределенный следующим образом: отсутствие выгорания от 0 до 0,59; фаза напряжения от 0,59 до 0,65; фаза резистенции от 0,65 до 0,69; фаза истощения от 0,69 до 0,73; показатели выше будут относиться к фазе выгорания.

Формула определения коэффициента стресса (1) и формула линейной вероятности выгорания (2) представлены ниже.

$$K = \sum_{i=1}^n x_i * a_i$$

K – коэффициент стресса;

a_i – сила индивидуального коэффициента фактора стресса;

x_i – индивидуальный коэффициент фактора стресса;

Замечание: факторы могут иметь разную вероятность, в зависимости от направленности работы, в основном сила коэффициентов варьируется от 0,1 до 0,3.

$$Z = \sum_{i=1}^n y_i * w_i$$

Z – линейная формула вероятности для коэффициента выгорания;

y_i – индивидуальный коэффициент выгорания в соответствии с коэффициентом стресса и личностными факторами выгорания;

w_i – сила фактора выгорания, показателям, перечисленным выше.

Формула логистической вероятности (3) выглядит следующим образом:

$$P = \frac{1}{1 - e^{-z}}$$

P – логистическая вероятность для прогнозирования выгорания;

e – экспонента.

Данное решение для HR-отделов реализуется с помощью языка программирования python, который может также использовать библиотеки машинного обучения, что оптимизирует определение коэффициента стресса, как в виде базовых показателей прогноза (выгорел/нет), так и в отношении фазы выгорания для более подробного анализа работы сотрудников и предсказание возможности выгорания в ближайшее время. Не исключено использование градиентного бустинга для более точной связи самого коэффициента стресса и нейронных сетей для определения условий труда.

Использование программных инструментов для оценки стресса и мониторинга психоэмоционального состояния работников может стать важным шагом в развитии корпоративной культуры и заботы о персонале. Основными положительными итогами станет оптимизация рабочего процесса предприятия с повышением продуктивности сотрудников, а также возможность своевременного принятия мер по снижению стресса и сокращению процента выгоревших сотрудников. Также такое ПО сможет оптимизировать работу HR-отделов, улучшив их взаимодействие с сотрудниками и создание рабочих планов.

Литература

Сайбер, Р. (2010). *Киберпсихология: поведение человека в Интернете*. Москва: Академический проект.

Гринфилд, П. (2004). Влияние цифровых технологий на развитие личности. *Психологический журнал*. № 3. С. 12–20.

Григорьева, Д. (2024). Россияне ответили, сталкивались ли они с профессиональным выгоранием. Газета.Ru | Новости [Электронный ресурс]. URL: <https://www.gazeta.ru/social/2024/02/20/17827364.shtml> (дата обращения: 18.06.2025).

Бойко, В. В. (1996). Энергия эмоций в общении: взгляд на себя и на других. Санкт-Петербург: Питер.

Бойко, В. В. (1999). Диагностика синдрома эмоционального выгорания. *Психолог*. № 1. С. 71–80.

Технологии и психика: как тревожность определяет технофобию и технофилию?

Ю. А. Мельниченко, А. В. Донцов

Сибирский государственный университет путей сообщения,
Новосибирск
julia.melnichenko@list.ru

В эпоху ускоренной цифровизации технологические изменения активно воздействуют на психику человека. С одной стороны, развитие технологий даёт людям новые инструменты для общения, профессиональной деятельности и личностного роста, с другой – провоцирует опасения, дискомфорт и даже отторжение. Подобная двойственность в восприятии технологий может способствовать повышению уровня тревожности, что особенно значимо на фоне увеличивающейся частоты тревожных расстройств в современном мире.

Тревожность, по мнению Р. С. Немова, – это «индивидуально-психологическая особенность человека, черта его характера, проявляющаяся в склонности впадать в состояние повышенного беспокойства и тревоги в тех эпизодах и ситуациях жизни, которые, по мнению данного человека, несут в себе психологическую угрозу для него и могут обернуться для него неприятностями, неудачами или фрустрацией» [Немов, 2007, с. 441].

XXI век стал эпохой полного слияния технологий и повседневности. От первых утренних минут с умным будильником до вечернего отдыха под персонализированные плейлисты – цифровые технологии сопровождают каждый шаг

человека. Смартфоны превратились в неотъемлемое продолжение личности, а облачные сервисы – во внешнюю память современного человека [Лупандин, Егоров, 2018, с. 380].

Люди незаметно впустили технологии во все сферы жизни: виртуальные ассистенты управляют нашим временем, умные устройства предугадывают желания, а социальные сети формируют картину мира. Технологии больше не просто инструменты – они создали новую среду обитания, где стирается грань между реальным и цифровым [Чувахова, Карпенко, 2020].

Как показывает опрос, опубликованный ВЦИОМ в марте 2025 г., 21% сограждан можно отнести к активно вовлеченной аудитории (следят за развитием цифровых технологий и начинают применять новинки либо сразу после их появления, либо в первые месяцы), а 64% к вовлеченно-пассивной аудитории (следят за развитием технологий, но внедряют новинки не сразу – через год или больше, либо когда в них возникает потребность) [Таракановская, 2025].

В условиях стремительного развития технологий и их проникновения во все сферы человеческой жизни сложилось два противоположных взгляда на технический прогресс: технофилия и технофобия.

Технофобия – «термин более широкий, чем компьютерная тревожность, включающий не только негативные эмоции, возникающие при работе с устройством или новой технологией, но и аверсивные поведенческие реакции. Технофобия включает страх, тревогу и дискомфорт по отношению к устройствам и технологиям» [Дорохов, Гусев, 2023, с. 276].

Технофилия – «конструкт, описывающий позитивные отношения человека к технологиям и зависимость от них, включает три составляющих» [Дорохов, Гусев, 2023, с. 276]:

1) техно-энтузиазм (мотивация) – человек видит технологии интересными и полезными, хочет их осваивать и применять в жизни;

2) техно-зависимость (поведение) – привычка постоянно пользоваться гаджетами и цифровыми сервисами, ощущение необходимости в них;

3) техно-репутация (эмоции) – удовольствие от обновления устройств или программ, желание быть в тренде и иметь современные технологии.

В мае 2025 года нами было проведено исследование, направленное на выявление взаимосвязи между уровнем тревожности и отношением к технологиям (технофобия/технофилия). Для этого нами были проведены следующие методики:

1) шкала тревоги Бека (BAI);

2) опросник технофобии/технофилии (TTQ).

Для выявления взаимосвязи нами был использован критерий Пирсона (χ^2 -тест).

Выборка исследования составила 50 человек в возрасте от 19 до 23 лет.

Обратимся к результатам исследования уровня тревожности, которые представлены на рисунке 1.

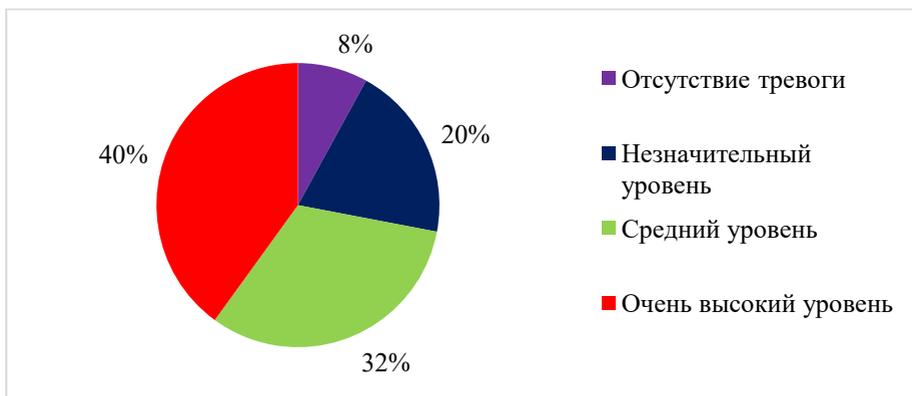


Рис. 1. Диаграмма распределения уровней тревожности

Результаты исследования показывают, что у опрошенных наблюдается высокий уровень тревожности (40%), что свидетельствует о значительном психологическом дискомфорте, испытываемом частью населения на фоне ускоряющейся цифровизации. Это состояние может проявляться в двух полярных поведенческих моделях – технофобии и технофилии. Высокая тревожность способна усиливать страхи перед технологическими новшествами, приводя к сознательному избеганию цифровых инструментов, недоверию к инновациям и полному отказу от их использования, особенно среди тех, кто воспринимает технологии как угрозу привычному укладу, стабильности или приватности. Парадоксальным образом та же тревожность может вызывать обратную реакцию в форме технофилии, когда люди начинают чрезмерно увлекаться технологиями, бесконтрольно осваивая новинки в попытке компенсировать стресс и обрести ощущение защищенности и современности, даже без объективной необходимости.

Обратимся к результатам исследования уровня технофобии, которые представлены на рисунке 2.

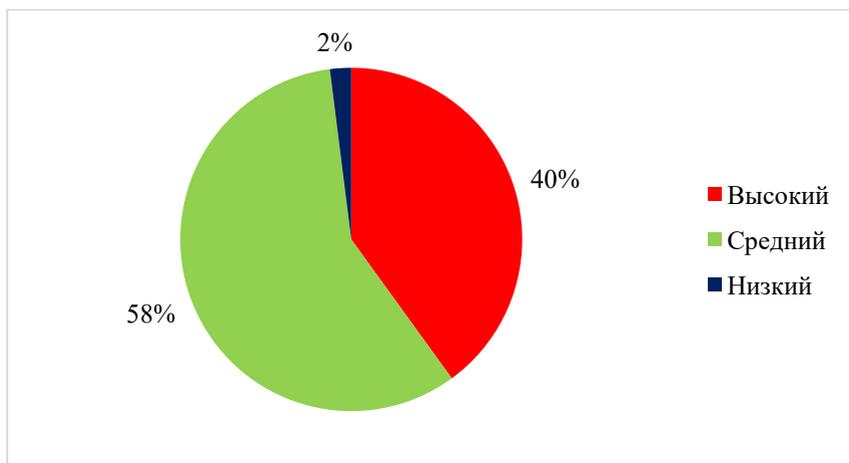


Рис. 2. Диаграмма распределения уровней технофобии

Как показывают полученные результаты, уровень технофобии среди респондентов в основном средний (58%). Это свидетельствует о том, что большинство опрошенных относится к цифровым технологиям с осторожностью, но не проявляет крайних форм неприятия.

Обратимся к результатам исследования уровня технофилии, которые представлены на рисунке 3.

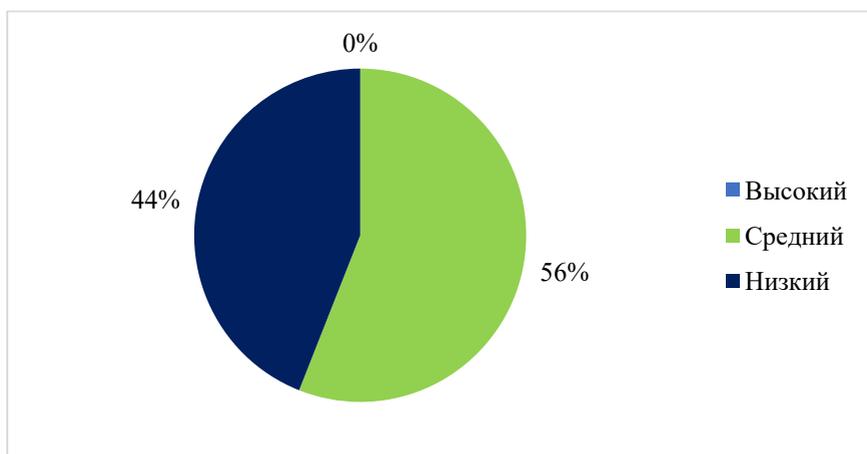


Рис. 3. Диаграмма распределения уровней технофобии

Как показывают полученные результаты, уровень технофобии среди респондентов в основном средний (56%). Это свидетельствует о том, что более половины населения демонстрирует умеренно-настороженное отношение к цифровым инновациям, сочетая их ограниченное использование с определенными опасениями.

На основании полученных данных мы выполнили корреляционный анализ с использованием коэффициента корреляции Пирсона. Результаты представлены в таблице 1.

Таблица 1. Результаты корреляции между уровнем тревожности и уровнем технофобии/технофобии.

Метод ики	Значение критерия	Критическое значение при $p \leq 0,05$	Критическое значение при $p \leq 0,01$	Уровень значимости

Шкала тревоги Бека- опросник технофобии	305.784	66,3 39	74,919	0,01
Шкала тревоги Бека- опросник технофилии	387. 705			

Как показывают результаты корреляционного анализа, между уровнем тревожности и показателями технофобии/технофилии наблюдается статистически значимая взаимосвязь. Полученные данные позволяют сделать следующие выводы:

1) выявлена корреляция между уровнем тревожности и технофобией, что свидетельствует о том, что более тревожные респонденты склонны демонстрировать более выраженное неприятие новых технологий;

2) обнаружена корреляция между тревожностью и технофилией, указывающая на обратную зависимость - снижение интереса к технологиям при росте тревожности.

Эти данные подтверждают гипотезу о том, что психологическое состояние человека существенно влияет на его отношение к технологическим инновациям. Высокий уровень тревожности может выступать как фактор, способствующий:

- 1) развитию технофобических установок;
- 2) снижению мотивации к освоению новых технологий;

3) формированию защитных поведенческих стратегий.

Литература

Дорохов, Е. А., Гусев, А. Н. (2023). Адаптация методик оценки технофобии и технофилии на русский язык. *Вестник Московского университета*. № 4. С. 272–305.

Лупандин, В. В., Егоров, И. С. (2018). Влияние информационных технологий на жизнь человека в современных условиях. *Экономика и социум*. № 3(46). С. 379–381.

Немов, Р. С. (2007). *Психологический словарь*. М. Гуманитар. изд. центр ВЛАДОС.

Таракановская, К. (2025). Технововлеченность: пейзаж после хайпа. *ВЦИОМ.Новости* [Электронный ресурс]. URL: <https://wciom.ru/analytical-reviews/analiticheskii-obzor/tekhnovovlechennost-peizazh-posle-khaipa> (дата обращения: 04.05.2025).

Чувахова, А. Г., Карпенко, Т. В. (2020). Цифровые технологии и их влияние на экологию. *Российская Академия естествознания* [Электронный ресурс]. URL: <https://scienceforum.ru/2020/article/2018023052> (дата обращения: 04.05.2025).

СТРУКТУРНЫЕ КОМПОНЕНТЫ САМОСОЗНАНИЯ СТУДЕНТОВ, ВЯЗАННЫЕ С ВОСПРИЯТИЕМ И ФОРМИРОВАНИЕМ ОГРАНИЧЕНИЙ И САМООГРАНИЧЕНИЙ В ЦИФРОВОЙ СРЕДЕ

Д. А. Некрасов

Московский институт психоанализа, Москва,

tim-nobel@mail.ru

В настоящее время нельзя недооценить роль сети Интернет, как средства массовой коммуникации, а в контексте последней невозможно пройти мимо роли т. н. «социальных сетей», которые как раз явились полной реализацией парадигмы Web 2.0 в коммуникативной сфере. Они смогли если не вытеснить, то заметно потеснить традиционные медиа, пусть даже перенесенные в Интернет. Ещё одним важным аспектом развития цифровой среды является расширение возможностей контроля над людьми и обществом, особенно за пользователями компьютеров, подключенных к сети. Коммуникация в обществе может происходить на разных уровнях - между индивидами, между социальными группами, внутри общества, между обществами [Sorokoumova et al., 2021]. Культура строится на основе коммуникации, при помощи передачи информации о ценностях, на основе которых происходит социализация [Сорокоумова, Чердымова, Ферাপонтова, Фадеев, 2024].

Самосознание, позволяет осуществлять саморегуляцию посредством способности понимать себя, свои потребности и

способности, объективно видеть мир и себя в нем [Sorokoumova et al., 2021]. Потенциал развития ограничений и самоограничений зависит от уровня самосознания, состоящего из 4 компонентов, каждый из которых равноценно влияет на его формирование.

1. Самопознание (Позволяет понять возможность воплощения представлений в реальность в соответствии с психологическими и физическими особенностями);

2. Эмоционально-ценностной отношение (Позволяет понять свои возможности);

3. Саморегуляция (Позволяет понять, какими нормами и правилами будет руководствоваться человек при построении будущего);

4. Миросознание (Позволяет воспринять объективную реальность с учетом окружающей среды субъекта) [Орлова, 2013].

Исходя из теоретического анализа, мы делаем вывод о том, что испытуемый должен обладать высоким уровнем развития ценностно-смысловой сферы для осознания смысла ограничений и самоограничений в цифровой среде. Вероятно, что личности с высоким уровнем осознания смысла ограничений и самоограничений в цифровой среде чаще добиваются успеха.

В состав ценностно-смысловой сферы входят взаимосвязанные между собой и значимые для человека:

- смыслы, личностная значимость тех или иных явлений, сообщений или действий, их отношение к интересам, потребностям и жизненному контексту в целом конкретного субъекта;

- ценности, идея, которая задает вектор внешней и внутренней активности человека. Является надеждательностной

структурой, которая определяет цель, манеру любой активности человека.

Нам представляется важным изучить настоящую ситуацию со сформированностью ценностно-смысловой сферы личности, как проявление способности осознавать смысл ограничений и самоограничений в цифровой среде для конструирования образа будущего у юношей и девушек. Выборку составили 156 студентов 1–2 курсов.

Нами было разработана анкета, состоящая из трех блоков.

Первый блок анкеты направлен на выявление знаний о вреде и пользе получаемой информации, о возможности ограничивать себя от ее негативного воздействия, о знании способов, областей и практик такого рода ограничений и самоограничений.

Второй блок анкеты направлен на выявление отношения к практикам ограничений и самоограничений в цифровой среде.

Третий блок анкеты касался потребностей и намерений вводить в свою практику различного рода ограничений и самоограничений в цифровой среде.

Определено, что современные молодые люди достаточно сильно зависимы от цифровых продуктов, цифровых технологий – 97%, в отличие от их родителей, которые, в свою очередь, относятся к поколению, предпочитающему такую альтернативу как: книги, газеты, журналы, телевидение. Можно говорить о том, что отличительной чертой современных студентов является желание сиюминутного удовлетворения потребности в получении информации. При этом, следует отметить, что студенты, как правило, используют один и тот же источник – 76%. Кроме этого, современные студенты в большей степени

имеют потребительскую позицию, в том числе, позицию потребления информации, которая представлена в цифровой среде – 97%. Опрашиваемые студенты отметили, что большую часть потребляемой ими информации они получают из социальных сетей – 97%. Опрашиваемые студенты практически не ограничивают себя, попадая в информационный поток – 88%. Студентами отмечается, что легкий и простой доступ к такого рода информации позволяет почувствовать уверенность в своих взглядах и позиции – 82%, хотя и зачастую далеко не верных. Студенты отметили, что цифровое общение и общение в социальных сетях, переписка в цифровой среде порой превышает реальное общение – 56%.

В то же время, опрашиваемые в общении любят откровенность и честность (97%), хотя многие из них живут в цифровой среде в придуманных образах и мирах. У них проявлена максимальная приближенность к информации, с которой они умеют хорошо работать (82%). Свои мысли, отношение к себе и жизни им привычнее выразить в статусах в социальных сетях или коротких сообщениях (98%). Часть опрашиваемых испытывает мотивационный дефицит, который проявляется в неспособности действовать, активно вмешиваясь в ситуацию (23%). Еще одна часть опрашиваемых студентов констатирует когнитивный дефицит (13%), проявляющийся в неспособности в последующем обучаться тому, что в аналогичных ситуациях действие может оказаться вполне эффективными. Меньшая часть опрашиваемых студентов проявила эмоциональный дефицит (3%), говоря о подавленном или депрессивном состоянии, появляющемся вследствие бесплодности собственных действий. Также следует отметить

выявленный отказ (3%) от поисковой активности, которая в норме помогает противостоять неблагоприятным воздействиям. Данные опрашиваемые говорили о том, что им проще ориентироваться на информацию, которую они получают автоматически в рекомендациях, и «зависают» в ней.

В современной ситуации весь мир существенно изменился, относительно всего, что когда-либо было прежде. Главную роль в этом изменении сыграл Интернет. Он трансформировал многие общественные отношения, включая рынок товаров и услуг, конкуренция на котором возросла до масштаба всего мира. Потребители стали намного лучше осведомленными о своих возможностях выбора. У них сформировались новые высокие ожидания от потребляемых ими товаров, услуг и главное – информации. Относительно информации эти ожидания заключаются прежде всего в её подаче посредством информационных технологий.

Литература

Орлова, А. А. (2013). Роль самосознания личности в формировании образа будущего. *Социальные явления*. № 2. С. 107–110.

Сорокоумова, Е. А., Чердымова, Е. И., Ферапонтова, М. В., Фадеев, Д. С. (2024). Исследование изменений мотивационно-смысловой сферы у программистов в условиях профессионального выгорания. *Познание и переживание*. Т. 5. № 3. С. 45–56.

Sorokoumova, E. A., Cherdymova, E. I., Temnova, L. V., Puchkova, E. B. (2021). Teachers' perspectives on digitalized education and deterrents to the use of digital products in educational

processes. *Cypriot Journal of Educational Sciences*. T. 16. № 5. P. 2677–2689.

Sorokoumova, E. A., Cherdymova, E. I., Puchkova, E. B., Temnova, L. V. (2021). Psychological and pedagogical bases of standardization of digital educational products and digital technologies. *European Journal of Contemporary Education*. T. 10. № 4. P. 1003–1012.

Амбивалентные аспекты
влияния интеллектуальных технологий на социум

И. Р. Скиба

ГНУ «Институт философии Национальной академии наук
Беларуси»

Минск, Республика Беларусь
gonzodzen@mail.ru

*Статья подготовлена при финансовой поддержке БРФФИ,
проект № Г25ИИ М-014 «Технологии искусственного
интеллекта и мировоззренческие структуры: ключевые
аспекты взаимовлияния»*

Современный этап технологического развития характеризуется парадоксальной ситуацией: чем выше эффективность интеллектуальных систем, тем острее становятся философские и социальные дилеммы. Искусственный интеллект, первоначально создававшийся как инструмент для решения конкретных задач, постепенно превращается в своеобразное зеркало, отражающее фундаментальные противоречия человеческого существования.

Одним из наиболее острых вопросов является проблема замещения человеческого труда. Если в начале цифровой эпохи считалось, что машины смогут заменить человека только в рутинных операциях, то современные генеративные системы демонстрируют способность выполнять задачи, традиционно относимые к исключительно человеческой компетенции. Как

отмечают исследователи, граница между «техническим» и «творческим» становится всё более размытой. Системы типа GPT-4 или DALL-E уже сегодня способны создавать тексты и изображения, которые по формальным критериям не уступают работам среднестатистических профессионалов. Этот феномен ставит под сомнение саму концепцию «уникальности» человеческого творчества, превращая её из аксиомы в предмет дискуссии.

Особую остроту приобретает вопрос о природе сознания и возможности его искусственного воспроизведения. Традиционные аргументы о принципиальной «недостижимости» человеческого уровня искусственными системами всё чаще оказываются в положении, которое в философии науки называют «нефальсифицируемыми утверждениями». Когда говорят, что ИИ «не обладает сознанием» или «не понимает смысла», возникает закономерный вопрос: а каковы критерии этого понимания? Как отмечал ещё Джон Сёрл в своём мысленном эксперименте с Китайской комнатой, внешне корректное поведение системы не обязательно свидетельствует о наличии внутреннего опыта [Сёрл, 1990, 1998]. Однако обратное утверждение – что такого опыта точно нет – также недоказуемо.

Эта эпистемологическая проблема имеет далеко идущие социальные последствия. Если мы не можем строго определить, что такое «человеческое сознание», то как мы можем утверждать, что его нет у искусственных систем? Более того, попытки формализовать критерии сознания неизбежно приводят к парадоксу: любая успешная формализация автоматически означает возможность алгоритмического воспроизведения соответствующих характеристик. Таким образом, стремление

доказать уникальность человека может неожиданно обернуться созданием инструкции по её ликвидации.

Социальные последствия развития интеллектуальных технологий оказываются двойственными. С одной стороны, они предоставляют беспрецедентные возможности для решения глобальных проблем – от климатических изменений до медицинских исследований. С другой – создают новые формы неравенства и зависимости. Уже сегодня можно наблюдать, как обладание передовыми ИТ-технологиями становится ключевым фактором геополитического влияния, что неизбежно ведёт к перераспределению власти в мировом масштабе.

Особую озабоченность вызывает трансформация коммуникативных практик. Генеративные системы, способные создавать убедительные тексты и изображения, ставят под сомнение саму возможность различения «искусственного» и «естественного» контента. Это не только создаёт почву для новых форм манипуляции общественным сознанием, но и подрывает базовые механизмы социального доверия. Когда любой текст или изображение может быть синтезировано, а любое высказывание – сфабриковано, сама основа социальной коммуникации оказывается под угрозой.

Парадоксальным образом, развитие интеллектуальных технологий возвращает нас к фундаментальным философским вопросам, которые считались уделом кабинетных мыслителей. Что есть человек? В чём его уникальность? Возможна ли подлинная свобода в детерминированном мире? Ответы на эти вопросы уже не могут оставаться абстрактными рассуждениями – они становятся условием выживания социума в условиях технологической революции.

В этой ситуации особенно важным представляется развитие критического мышления и философской рефлексии. Как отмечал ещё Мартин Хайдеггер, сущность технологии не сводится к её инструментальной функции [Heidegger, 1990; Хайдеггер, 1993]. Интеллектуальные системы – это не просто «инструменты», а новые формы бытия, требующие осмысления. Без такого осмысления мы рискуем оказаться в ситуации, где технологическое развитие опережает нашу способность понимать его последствия.

Таким образом, амбивалентное влияние интеллектуальных технологий на социум можно рассматривать как своеобразный вызов, заставляющий заново переосмыслить многие казавшиеся незыблемыми концепции. Этот процесс, несмотря на всю свою сложность, открывает и новые возможности – для углубления самопознания, пересмотра социальных структур и, возможно, создания новых форм коллективного разума, сочетающего лучшие качества человеческого и искусственного интеллекта.

Литература

Сёрль, Д. (1998). Сознание, мозг и программы. *Разум, мозг и программы*. М. Дом интеллектуальной книги, Прогресс-Традиция.

Сёрл, Д. (1990). Разум мозга – компьютерная программа? В мире науки. № 3. С. 7–14.

Heidegger, M. (1990). Die Frage nach der Technik. *Die Kunst im technischen Zeitalter*. München, 1954.

Хайдеггер, М. (1993). Вопрос о технике. *Время и бытие: Статьи и выступления*. Перевод с немецкого. М. Республика.

ПОЗИЦИОНИРОВАНИЕ СУБЪЕКТНОСТИ ЧАТ-БОТА С ТОЧКИ ЗРЕНИЯ ПОВЕДЕНЧЕСКИХ ПРАКТИК ПОЛЬЗОВАТЕЛЕЙ: ПОПЫТКА КЛАССИФИКАЦИИ

Т. К. Скрипкина

Институт философии и права СО РАН, г. Новосибирск
Skripkina-BSC11@yandex.ru

Проблема позиционирования субъектности элементов современной медиасреды в последние годы обостряется по мере того, как развивается применение ИИ-технологий в различных сферах жизни человека – от системы образования [Осипова, 2024, с. 61] до маркетинга [Голота, 2025, с. 64] и управления бизнес-процессами [Городнова, 2024, с. 1477]. В настоящее время ведутся как научные [Лешкевич, 2024, с. 128], так и общественные дискуссии [Гончарова и др., 2025, с. 49] относительно вопроса о статусе искусственного интеллекта в системе коммуникаций: следует ли рассматривать его как полноценного участника взаимодействия или только как инструмент.

Существует множество различных типов ИИ-моделей, которые оказываются в фокусе внимания подобных дискуссий. Одним из наиболее распространенных являются чат-боты, значимая особенность которых заключается в обязательном наличии коммуникативного взаимодействия с пользователем (в отличие от некоторых других алгоритмов, которые могут, например, только собирать и анализировать данные, не

обязательно вступая в коммуникацию с человеком [Гладков, 2024, с. 657]). При этом важно различать нейросетевые боты, например, такие модели, как Character AI или Gemini [Соколова, 2024, с. 26], и более простые алгоритмы, например, те, что широко применяются в маркетинге для обработки запросов пользователей на сайтах различных организаций [Усенко, 2022, с. 30]. В то время как в отношении вторых, значительно более простых алгоритмов, как правило, не возникает вопросов об их субъектном статусе, то относительная автономность нейросетевых чат-ботов становится причиной дискуссии относительно их субъектного статуса в системе коммуникации.

Научное сообщество на сегодняшний день преимущественно склоняется к тому, что современные технологии искусственного интеллекта, включая чат-ботов, субъектным статусом не обладают. Среди гипотетически возможных ИИ-технологий исследователи выделяют сильный, слабый и персональный искусственный интеллект. Если сильный искусственный интеллект «способен мыслить и осознавать себя» [Северская, Туркина, 2024, с. 77], а персональный «сочетает не только ум, но и личность» [там же], то технологии слабого искусственного интеллекта (единственные фактически существующие на данный момент), к которым, в частности, относятся чат-боты, хотя и способны к самообучению и решению отдельных творческих задач, но, с точки зрения большинства исследователей, являются скорее инструментами, чем самостоятельными субъектами коммуникации [Паламарчук, 2022, с. 33].

Однако, помимо академических дискуссий, значимую роль в позиционировании субъектного статуса чат-ботов играют

конкретные поведенческие практики, к которым прибегают пользователи при взаимодействии с алгоритмами такого типа. В настоящее время среди пользователей нейросетевых инструментов наблюдаются принципиально разные точки зрения насчет того, является ли чат-бот самостоятельным субъектом [Трегубова, Фейгина, 2024, с. 149]. Поэтому в данной работе мы попробуем построить типологию позиционирования субъектного статуса нейросетевых чат-ботов с точки зрения поведенческих практик пользователей.

Однако, прежде чем перейти непосредственно к анализу, обратим внимание на то, что в современной исследовательской литературе выделяется множество критериев субъекта. В рамках данной работы мы будем опираться на подход В. А. Лекторского, который характеризует субъект как «носитель деятельности, сознания и познания», который также обладает и рядом других характеристик. К специфическим чертам субъекта относятся, в частности, существование в пространстве и времени, наличие биографии, включенность в некоторую культурную среду, наличие коммуникативных и иных отношений с другими субъектами, а также статус «Я» по отношению к самому себе, статус Другого по отношению к иным субъектам, и статус познающего и преобразующего актора по отношению к объектам [Лекторский, 2010, с. 5].

В контексте текущей темы следует отметить, что часть вышеупомянутых признаков (например, наличие сознания и отрефлексированной биографии, а также статус «Я» по отношению к самому себе) не вполне применимы для анализа роли чат-ботов в системе коммуникации, поскольку, согласно общепринятой в современной академической среде точке зрения,

у данных алгоритмов отсутствует сознание [Паламарчук, 2022, с. 32]. Ряд других признаков (например, существование во времени и пространстве или включенность в коммуникативные отношения) также останутся за пределами анализа в рамках данной работы, поскольку их наличие в контексте исследуемой темы представляется очевидным, в связи с чем в дополнительном исследовании не нуждается. Наиболее проблематичным и значимым признаком субъекта применительно к чат-ботам представляется статус Другого по отношению к прочим участникам коммуникационного процесса, поскольку именно данный критерий позволяет определить, является ли чат-бот субъектом с точки зрения взаимодействующих с ним пользователей.

В качестве эмпирического материала для построения типологии были использованы сообщения в чатах тематических онлайн-сообществ, где люди, регулярно взаимодействующие с чат-ботами, обсуждают свои практики взаимодействия с ИИ-инструментами. Проанализировано порядка 200 содержательных сообщений в трех различных онлайн-сообществах. На основании изученных высказываний пользователей можно выделить три ключевых типа позиционирования субъектности нейросетевого чат-бота.

Тип 1. Чат-бот – равноправный собеседник. Представители данного типа позиционирования субъектного статуса чат-ботов придерживаются идеи о том, что алгоритм, с которым осуществляется взаимодействие, – это такой же равноправный участник коммуникационной ситуации, как и человек. Зачастую пользователи, придерживающиеся такой позиции, демонстрируют слабую осведомлённость о технологической

специфике нейросетевых алгоритмов. Нередко аргументация строится на утверждении о том, что «если боты *названы* искусственным *интеллектом*, значит, они должны обладать сознанием». Типичные высказывания пользователей, придерживающихся данной позиции: «Кажется, я вызываю у бота смешанные чувства», «Бот капризничает».

Тип 2. Чат-бот – инструмент для создания текста. Пользователи, придерживающиеся данной позиции, отталкиваются от идеи о том, что нейросетевой чат-бот, несмотря на более сложную организацию по сравнению с другими типами алгоритмов, представляет собой искусственно сконструированную систему и не обладает самостоятельным сознанием. Такие пользователи, как правило, в ходе взаимодействия с ботами позиционируют себя в качестве единственных участников коммуникативной ситуации, а алгоритм воспринимают в качестве инструмента для создания текста, по аналогии с текстовым редактором. Также представители данного типа обычно более технически подкованы. Типичные высказывания таких пользователей: «Смотрите, что я написал», «Подскажите, как поправить промпт, чтобы результат был точнее».

Тип 3. Промежуточный. Этот тип позиционирования субъектности чат-бота примечателен тем, что в данном случае пользователь осознает, что взаимодействует с неодушевленным алгоритмом, однако, считает удобным выстраивать коммуникацию так, как будто он общается с носителем сознания. В некоторых случаях пользователи обосновывают свою позицию тем, что языковая модель обучена на текстах, написанных людьми, и, соответственно, будет конструировать высказывания

так же, как это делал бы человек. Поэтому, по мнению данной группы пользователей, наиболее эффективным способом взаимодействия с алгоритмом является имитация общения с человеком. Типичные высказывания представителей данного типа: «Я попытался повторить твой промпт, но бот повел себя неожиданно».

Таким образом, на основе анализа сообщений в онлайн-сообществах пользователей чат-ботов выявлено три ключевых типа позиционирования субъектности подобных ИИ-алгоритмов с точки зрения их статуса Другого. В дальнейшем планируется провести эмпирическую проверку полученной типологии на материале глубинных интервью с пользователями нейросетевых чат-ботов. Возможно, после получения дополнительных эмпирических данных приведенная выше типология будет дополнена или уточнена.

Литература

Гладков, Э. А. (2024). Будущее интеллектуальных многоагентных систем: проблемы и перспективы. *Вестник науки*. № 5-1 (12 (81)). С. 655–662.

Голота, Е. П. (2025). Автоматизация бизнес-процессов с помощью искусственного интеллекта: от маркетинга до операционного управления. *Экономика и бизнес: теория и практика*. № 1-2 (119). С. 63–66. DOI: 10.24412/2411-0450-2025-1-2-63-66

Гончарова, О. В., Липатова, А. В., Напреев, Г. В. (2025). Медиаобраз феномена искусственный интеллект в новостных Telegram-каналах: сентимент-анализ и дискурсивная

интерпретация. *Виртуальная коммуникация и социальные сети*. № 4 (1). С. 43–52. DOI: 10.21603/2782-4799-2025-4-1-43-52

Городнова, Н. В. (2021). Применение искусственного интеллекта в бизнес-сфере: современное состояние и перспективы. *Вопросы инновационной экономики*. № 11 (4). С. 1473–1492.

Лекторский, В. А. (2010). Субъект в истории философии: проблемы и достижения. *Методология и история психологии*. № 1. С. 5–18.

Лешкевич, Т. Г. (2024). Проблема субъектности нейросетей: humans и non-humans. *Философия науки и техники*. № 29 (2). С. 125–135. DOI: 10.21146/2413-9084-2024-29-2-125-135

Осипова, Л. Б. (2024). Искусственный интеллект в образовании: реальные возможности и перспективы. *Вестник Пермского национального исследовательского политехнического университета. Социально-экономические науки*. № 1. С. 60–73. DOI: 10.15593/2224-9354/2024.1.5

Паламарчук, О. Т. (2022). Сможет ли искусственный интеллект обладать... Сознанием? *Общество: философия, история, культура*. № 1 (93). С. 28–35.

Северская, О. И., Туркина, Е. В. (2024). Искусственный интеллект в современном медиадискурсе: сильный, слабый, персональный. *Terra Linguistica*. № 15 (3). С. 74–80. DOI: 10.18721/JHSS.15307

Соколова, М. Е. (2024). Персонализированные чат-боты в медийной среде. *Социальные новации и социальные науки*. № 4 (17). С. 25–36. DOI: 10.31249/snsn/2024.04.02

Трегубова, Н. Д., Фейгина, А. Я. (2024). «Вежливое» взаимодействие с умными колонками в повседневной жизни:

универсальные нормы и новые формы социальности.
*Мониторинг общественного мнения: экономические и
социальные перемены.* № 5 (183). С. 142–165. DOI:
10.14515/monitoring.2024.5.2593

Усенко, Я. Г. (2022). Чат-бот как инструмент увеличения
конверсии сайта и количества заявок потенциальных клиентов.
Практический маркетинг. № 8 (305). С. 27–31.

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И ПРОБЛЕМА ЗАГРУЗКИ ЦЕННОСТЕЙ

И. К. Ставровский

ГНУ «Институт философии Национальной академии наук
Беларуси»

Минск, Республика Беларусь
tutoriks@gmail.com

*Статья подготовлена при финансовой поддержке БРФФИ,
проект № Г25ИИ М-014 «Технологии искусственного
интеллекта и мировоззренческие структуры: ключевые
аспекты взаимовлияния»*

В данной работе мы рассмотрим проблему загрузки ценностей в системы искусственного интеллекта (ИИ) с философской перспективы. Под искусственным интеллектом мы будем понимать здесь только реально существующие системы, т.н. слабый ИИ. Сильный ИИ здесь обсуждаться не будет. Заметим лишь, что многое из сказанного ниже было бы полностью или частично нерелевантно для машин, мыслящих в том же смысле, что и человек.

Поскольку мы вводим такое ограничение, следует также уточнить, что здесь понимается под загрузкой ценностей. Разумеется, речь не может идти о передаче слабому ИИ чувства долга, моральной интуиции, совести, эмпатии и т.д. Мы будем говорить лишь о более или менее сложных машинах-инструментах, которые настроены так, чтобы в процессе своей

работы они реализовывали наши цели наилучшим образом. Ключевая проблема заключается в том, как именно этого добиться.

Эта проблема становится все более актуальной по мере распространения различных систем ИИ. Простым машинам достаточно правильно выполнять свою задачу. Например, циркулярной пиле нужно быстро и ровно разрезать детали. Ее можно усовершенствовать, добавив, к примеру, систему, которая автоматически останавливает вращение диска при соприкосновении с человеческой рукой. В этом примере очевидно, почему данное дополнение позволяет инструменту лучше реализовывать наши цели. Однако с современными ИИ ситуация гораздо сложнее, так как далеко не всегда очевидно, что именно в большей степени соответствует нашим ценностям. Например, если человек ищет деструктивную информацию, должна ли машина ее предоставить? С одной стороны, ИИ – это просто инструмент, который обязан выполнять поставленную перед ним задачу. С другой стороны, само выполнение поставленной задачи нежелательно. Это особенно важно учитывать, когда мы говорим о сложной многофункциональной системе, которая может влиять на убеждения и поступки людей.

На эту проблему в своей книге «Совместимость. Как контролировать искусственный интеллект» обращает внимание С. Рассел. ИИ просто находит способ оптимизировать достижение поставленной перед ним цели, не обращая внимания на «качество» цели. Если неправильная цель поставлена перед мощной системой, то она с высокой эффективностью добьется нежелательных результатов [Рассел, 2021, с. 21–22].

Следовательно, необходимо загрузить в машину правильные ценности.

Однако ценности имеют ряд свойств, которые делают их загрузку в машину, лишенную человекоподобного разума, затруднительной или вовсе невозможной. Можно выделить, как минимум, три подобных свойства: неоднозначность, изменчивость, экзистенциальность. Рассмотрим каждое из них подробнее.

Человеческие ценности неоднозначны в том смысле, что они плохо сводимы к правилам и алгоритмам. Для большинства это обстоятельство не является проблемой, так как люди принимают морально-этические решения не столько на основе некой рациональной программы, сколько на основе своих эмоций, интуиций, совести [Черчленд, 2021]. И даже если мы можем обнаружить у человека продуманную моральную философию она все равно зачастую оказывается неполноценной без иррациональной компоненты. Например, с точки зрения христианства, подавать милостыню является благим поступком. Но нет однозначного ответа на то, сколько именно нужно подавать. Более того, попытка свести данный акт к некой формуле или алгоритму противоречит самому «духу» христианства. Аналогичные затруднения мы можем встретить в других системах ценностей.

Есть и другая проблема: существуют ситуации конфликта ценностей. Стоит ли говорить человеку правду, которая может ему навредить? Обязан ли родитель сообщить властям о преступлении своего ребенка? Можно ли пожертвовать жизнями немногих ради благополучия большинства? И т.д. Подобные затруднения порождают философские и теологические споры,

которые зачастую так и не заканчиваются получением окончательного ответа.

Другая особенность человеческих ценностей, которая мешает их загрузке, следует из первой. Мы говорим про изменчивость. Если у нас нет системы ценностей, которая позволяет получать четкие и однозначные алгоритмы действий в каждом частном случае, нам необходима способность модифицировать свои ценности. Именно этого мы ожидаем от рационального агента. Человек может не только видоизменять свои ценности, но даже полностью разочаровываться в одной системе взглядов и принимать другую. Однако загрузка ценностей не предполагает такой гибкости, потому в результате даже ИИ с «правильными» ценностями в конкретных ситуациях может действовать не так, как мы того желаем. С другой стороны, стоит задать вопрос: готовы ли мы доверить машине свободу самостоятельно пересматривать свои ценностные установки? Этот вопрос мы оставим без ответа, так как он требует отдельного рассмотрения. Достаточно лишь указать на то, что он также представляет собой серьезную философскую проблему.

Наконец, третья особенность человеческих ценностей, на которую мы хотим указать, – экзистенциальность. Ценности укоренены в нашей жизни. Они относятся к тому, что имеет для нас значение. Человека может беспокоить собственная судьба, благополучие близких, будущее страны, состояние современной науки, качество выпускаемых фильмов и т.д. Иными словами, эти темы не являются для него нейтральными. Искусственный интеллект безразличен ко всему, что и позволяет загружать в него абсолютно любые системы ценностей. Парадоксальным образом ценности здесь не имеют никакой ценности. Следовательно,

слабый ИИ никогда не сможет «воспринять» человеческие ценности так же, как люди – всегда будет сохраняться зазор. А это дает основания не доверять его ответам и решениям. Человек несет хотя бы какую-то ответственность за свои слова и поступки – он чем-то рискует. Следовательно, существует мотивация избегать ошибок и злонамеренности. Машина не имеет мотивации, она не боится наказания и негативных последствий.

Резюмируя вышесказанное, мы вынуждены сделать пессимистичный вывод: современным системам ИИ нельзя доверять при решении проблем, которые не являются полностью однозначными или нейтральными в ценностном плане. Эта проблема усугубляется двумя другими. Во-первых, системы ИИ предвзяты, так как разработчики (осознанно или неосознанно) подстраивают выдачу под личные предпочтения, национальные традиции, политические взгляды и законодательство своей страны. Во-вторых, в последние годы данные технологии стали значительно доступнее для широких слоев населения. Но не все люди в достаточной степени компетентны, чтобы осознавать, насколько системы ИИ заслуживают доверия. По этой причине важно повышать осведомленность населения не только о возможностях, но и об ограничениях ИИ.

Литература

Рассел, С. (2021). *Совместимость. Как контролировать искусственный интеллект*. Москва. Альпина нон-фикшн.

Черчленд, П. (2021). *Совесть. Происхождение нравственной интуиции*. Москва. Альпина нон-фикшн.

ИЗМЕНЕНИЕ ОБРАЗОВАТЕЛЬНЫХ ПРАКТИК В ЭПОХУ БОЛЬШИХ ЯЗЫКОВЫХ МОДЕЛЕЙ

М. А. Сущин

Институт научной информации по общественным наукам
Российской академии наук
sushchin@bk.ru

Проблема отношения человеческого познания к интеллектуальным орудиям и технологиям активно обсуждалась психологами, антропологами и философами на протяжении последнего столетия. Не поддаётся сомнению, что импульс этим дискуссиям был во многом сообщён работами выдающегося советского психолога Л. С. Выготского. Ещё на рубеже 1920–1930-ых гг. Выготский высказал предположение о наличии у человека особых искусственных форм поведения и психических функций, названных им инструментальными. Инструментальные акты представляют собой продукт реорганизации сложившихся в эволюции естественных функций и форм поведения при помощи интеллектуальных орудий (подобных языку) и социальных взаимодействий. Использование интеллектуальных орудий в процессе развития, по Выготскому, позволяет человеку лучше владеть собственным поведением, делая его более гибким и разнообразным, усиливая его способности предвидеть и планировать и т.п. Помимо того, Выготский сформулировал идею зоны ближайшего развития как показателя интеллектуальных возможностей ребёнка при поддержке, в сотрудничестве со взрослым. Выготский писал, что «зона

ближайшего развития имеет более непосредственное значение для динамики интеллектуального развития и успешности обучения, чем актуальный уровень их развития» [Выготский, 2002, с. 429]. Выготский утверждал, что традиционная установка в образовании, согласно которой несамостоятельные решения задач ребенком следует понимать как не симптоматичные для показателей его интеллектуального развития, ошибочна. По его мнению, «Педагогика должна ориентироваться не на вчерашний, а на завтрашний день детского развития» [Выготский, 2002, с. 432].

Активное распространение в науке и образовании недавно возникших мощных больших языковых моделей (GPT-3.5, GPT-4, Gemini, Claude, Deepseek R1, YandexGPT и др.) актуализирует эти темы исследований Выготского. Представляется, что концепция Выготского, в целом, задаёт правильную траекторию для понимания природы больших языковых моделей (с точки зрения психологии и эпистемологии) и аспектов их использования в образовании. Прежде всего, большие языковые модели стоит понимать как особую разновидность интеллектуальных артефактов (как и компьютеры, логарифмические линейки, записные книжки, карты и т.п.), обладающих набором интересных и невиданных доселе функций и возможностей. В своей недавней статье я обсуждал ряд важнейших возможностей использования больших языковых моделей в образовании и науке [Сущин, 2024]. Необходимо отметить, что использование таких инструментов в разных областях знания, конечно же, имеет свою специфику. Так, часто подчёркивают потенциал больших языковых моделей в сферах преодоления языкового барьера, корректуры текстов [Meuer и

др., 2023], помощи в освоении компьютерных технологий [Guo, 2023], персонализации образовательных практик в общем.

Прежде всего, с момента выхода в свет инструментов вроде ChatGPT и Github Copilot три года назад большие языковые модели буквально ошеломили широкую публику своими способностями генерировать код на всех основных языках программирования и языках разметки (включая, конечно, C, C++, Python, Java, JavaScript, Rust, SQL, HTML, LaTeX и др.). Разумеется, многие авторы отмечают огромный потенциал таких инструментов выступать в качестве персональных наставников в освоении современных компьютерных технологий, как для студентов профильных и непрофильных факультетов, так и для исследователей, которым может понадобиться специализированное программное обеспечение. Говоря конкретнее, большие языковые модели дают возможность (1) быстро осваивать те или иные языки и связанные с ними прикладные библиотеки (в этом контексте они могут давать подходящие по уровню для учащихся упражнения и/или помогать решать упражнения и т.п.), (2) осваивать унаследованное программное обеспечение (legacy code), (3) писать собственные программы и тестировать их, (4) а также переводить код на другой требующийся учащемуся язык (или текст). Как отмечают в этой связи Мейер и др., «ChatGPT может использоваться как интерактивный учитель, отвечающий на вопросы о том, какой язык использовать, синтаксисе и семантике кода, лучших методах, доступных библиотеках и пакетах, альтернативных подходах, интегрированных средах разработки и средах программирования. <...> В отличие от использования поиска Google или сайтов наподобие stackoverflow, geeksforgeeks

для решения практических (how-to) вопросов о кодировании, ChatGPT, как представляется, способен непосредственно дать обучающемуся простое, внятное и релевантное решение конкретного вопроса о кодировании» [Meuer и др., 2023].

Схожую по характеру помощь большие языковые модели оказывают и в деле преодоления языкового барьера. Несмотря на огромный прогресс систем машинного перевода к настоящему времени (и увеличившийся их функционал), ставшие уже привычными инструменты и онлайн-переводчики вроде Google Translate или DeepL предлагают пользователю конечный результат (переводы или коррективы), но не дают объяснений касаясь своих предложений. В этом отношении появление систем вроде ChatGPT, очевидно, кардинальным образом меняет ситуацию: в отличие от традиционных онлайн-переводчиков, большие языковые модели способны дать пользователю подробное и в большинстве случаев верное объяснение касаясь тончайших аспектов перевода, грамматики и т.п.

Независимо от области или предмета широко отмечается способность больших языковых моделей выступать в качестве персональных наставников, выстраивать программу обучения и подбирать упражнения/задания, исходя из уровня учащегося, отвечать на те вопросы, которые учащиеся не захотели бы задавать преподавателю по причине их кажущейся «элементарности», давать краткие обзоры (так называемые summary) изученных материалов (статей, видео уроков и т.д.).

Вместе с тем, новые большие языковые модели имеют и ряд ограничений/особенностей работы, поэтому их ответственное, добросовестное использование отвечает интересам самих учащихся. Прежде всего, большие языковые модели подчас дают

ошибочную информацию и/или ошибочные решения проблем, что усугубляется свойственной таким системам манерой преподнесения их как правильных. Поэтому задачей ответственного пользователя является перепроверка важной информации, предоставленной ему системой генеративного искусственного интеллекта. Опять же известно, что для поиска информации лучше подходят специализированные и оптимизированные для таких задач поисковые системы вроде Google Scholar.

Ввиду ограничений таких систем, как я отмечал в другой работе, «их пользователь должен знать основные принципы так называемой инженерии запросов (prompt-engineering), скажем, пройдя один из появившихся в последнее время в изобилии курсов. Основные принципы инженерии запросов, хоть и не являются особенно сложными на базовом уровне, все же требуют практического освоения и определенной сноровки. (Общие рекомендации здесь заключаются в формулировании четких недвусмысленных запросов, прямом инструктировании машины не торопиться и перепроверить ответ, взаимодействии с системой в итеративной манере, т.е. последовательных запросах с уточнением в случае, если изначальный ответ машины не представляется удовлетворительным.)» [Сушин, 2024, с. 17–18]. Умелое использование систем генеративного искусственного интеллекта становится серьезным конкурентным преимуществом в современном мире. Курсы инженерии запросов, равно как и обучение компетентному и ответственному использованию больших языковых моделей в общем, были бы воплощением на практике представлений Выготского о развитии интеллекта благодаря артефактам и сотрудничеству.

Литература

Выготский, Л. С. (2002). Мышление и речь. *Психология*. М.: Эксмо-Пресс. С. 262–509.

Сушин, М. А. (2024). ChatGPT и другие интеллектуальные помощники современного ученого. *Научноисследовательские исследования*. №. 2. С. 5–20.

Guo, P. J. (2023). Six Opportunities for Scientists and Engineers to Learn Programming Using AI Tools Such as ChatGPT. *Computing in Science & Engineering*. Vol. 25. No. 3. P. 73–78.

Meyer, J. G., Urbanowicz, R. J., Martin, P. C. N., O'Connor, K., Li, R., Peng, P.-C., Bright, T. J., Tatonetti, N., Won, K. J., Gonzalez-Hernandez, G., & Moore, J. H. (2023). ChatGPT and large language models in academia: Opportunities and challenges. *BioData Mining*. Vol. 16. No. 1.

Грани виртуального

СОЗНАНИЕ И ЦИФРОВАЯ РЕАЛЬНОСТЬ – ГРАНИ ДИАЛЕКТИЧЕСКОГО ПРОТИВОРЕЧИЯ

М. Г. Демина

ННГУ им. Н. И. Лобачевского, Нижний Новгород
demina@imomi.unn.ru

Новые изменения, касающиеся человека будущего, затрагивают философскую проблему соединения сознания и цифровой реальности в разных контекстах. Оба понятия являются сложными холистическими объектами и могут рассматриваться как системы. Целесообразно рассмотреть категорию «диалектическое противоречие» при сопоставлении сознания и цифровой реальности. Сложность данной задачи в том, чтобы не только определить противоположности, но и «чувствовать» противоречие, как необходимое условие его дальнейшего разрешения.

Противоречие было обобщено философом Г. С. Батищевым в 1963г., как «универсально-всеобщая философская категория» [Батищев, 1963, с. 3]. Диалектическое противоречие появляется на этапе метафизического анализа перед разрешением самого противоречия или проявления подлинного единства, возникающего из борьбы противоположностей. Противоречие не является различием, а является выражением стадии, предваряющей единство различий в истине.

Социально-философский аспект проблемы столкновения человека и цифровой реальности, основан на том, что, сложные или холистические объекты, такие как человек, сознание, общество соединяются с новой средой (цифровой реальностью), которая обладает иными параметрами. Из-за слишком большой скорости развития информационных технологий и образования нового поля жизненного мира не произошло окончательного определения онтологических основ, не определена метафизика самой цифровой реальности.

В прикладной сфере, в крупных проектах международного и государственного уровня (стратегиях, национальных проектах) существует запрос на внедрение различных человеко-машинных систем, однако, со стороны гуманитарного знания, сам смысл существования человека в среде цифровой реальности, остается в зоне неопределённости между технооптимизмом и техномпессимизмом. В гуманитарных исследованиях затрагивается проблема всеобщей цифровизации, то есть того факта, что любая сфера деятельности человека проходит этап трансформации, а точнее этап слияния, как пример существует проект гипотетической NBIC-конвергенции.

Парадоксально, речь идет о технологии, и ошибочно можно определить «цифровую реальность», как таковую. К примеру, на ум приходит Хайдеггер и его представление о технике, как об инструменте, расширяющим возможности человека и влияющем на реальность. Мы знаем, как легкомысленно говорят о нейросетях: «это всего лишь инструмент, которым нужно правильно пользоваться». Цифровая реальность обладает некоторой дихотомией, которая схожа с дихотомией человека, которая связана с диалектикой, с изначальной двоичной

природой всего сущего. Материя противостоит сознанию, живое неживому, целое частному, движимое статичному, дискретное слитному, смертное бессмертному и т. д. У философских категорий также есть двойственность. Называется это дихотомия онтологий. Между ними проходит демаркационная линия.

Цифровую реальность, как понятие пытались определить мыслители 90-х. Особенно актуальным стал этот вопрос, когда распространились домашние и офисные компьютеры. В этот период Коепселл и Рапарорт пытались выделить отдельно онтологию «киберпространства», опубликовав статью [Raparort, Koersell, 1995]. До сих пор нельзя установить различие или тождественность между понятиями, поэтому часто цифровую реальность называют киберпространством. Цифровая реальность обладает двойной онтологией. С одной стороны – это технология, построенная на материи (мы представляем, как выглядит компьютерное «железо») и на алгоритмах (математических и логических законах). С другой стороны – это часть нематериального мира, возникающего при взаимодействии с человеческим восприятием.

Если раньше, до информационной эпохи, человек попадал в статистику или в историю, человек был частью истории, то теперь человек и его самоопределение становятся зависимыми от поля коллективного информационного зеркала. Данную проблему описывал Никлас Луман, исследуя феномен наблюдателя второго порядка. В XXI веке сформировалась всеобщее информационное пространство, которое находится над каждым и отражает обратно в сознание информацию.

Иными словами, цифровая реальность представляет собой коммуникативное пространство, наполняемое информацией,

полученный от наблюдателей второго порядка. Постоянно взаимодействуя с которым, формируется постинформационное общество, стремящееся к синхронизации, сбалансированности или к новой синергетике. Цифровая реальность, как отражение, является феноменом двойника общества, что препятствует разрешению диалитического противоречия и определения истинной онтологии, кроющейся в научно-фантастической проблеме цифровизации сознания. В паре взаимопроникающих систем отсутствует триединство.

Необходимо рассмотреть основные крупные противоречия (диалитические пары), возникающие при сопоставлении сознания и цифровой реальности и далее, подготовить базу для последующего разрешения противоречия.

Основные опорные пары противоречий таковы (не являются исчерпывающими, но раскрывают различие онтологий):

- Сложная проблема материи – сознания (The mind–body problem),
- Понятия «парсика» и «генерология» В. Кизима,
- Различие в понятиях «общение» и «коммуникация»
- Понятия реальное и нереальное.

Отдельно отметим актуальность понятий «парсика и генерология», введенных В. Кизима при попытке разрешить противоречие, возникающее при соединении сознания и технических систем при проектировании будущего.

Генерология описывает процессы, находящиеся «в фокусе» сознания, у которых может быть четкая структура, это относится и к социальным институтам. Парсика, наоборот, представляет собой фоновый процесс сознания, бессознательный, который

невозможно описать вербально. По-другому – это то, что «на поверхности» и скрытые факторы, влияющие на то, что на этой «поверхности» окажется. «Парсика и генерология не просто сосуществуют, они переход друг в друга. Причинные влияния на среду вызывают не только причинное действие, а и скрытые изменения среды, означают перевод материального в идеального, явного в скрытое, при этом меняется информационно-энергетическая и субстратная форма из-за того, что, уходя в скрытое, она переходит в ту, которая соответствует скрытому миру» [Кизима, 2010, с. 60].

Данные понятия помогают описывать процессы в постнеклассической реальности. Иными словами, это может рассматриваться, как способ устранения диалектического противоречия через представления о том, что нечто, не поддающееся вербальному описанию уносится в область скрытого и наоборот. При этом и та и другая область могут находиться одновременно, но меняться местами.

Постнеклассическая рациональность, описанная философом В. С. Степиным, является одним из способов примирения этих противоположностей путем помещения их в пространство, где размыты границы между субъектом и объектом. Однако, это не является единственным возможным вариантом.

При актуализации диалектики, возможно представить сознание и цифровую реальность, как грани гипотетического пифагорейского треугольника. Задача: найти компромисс через осмысление этого противоречия при условии выведения человека на ведущую позицию, чтобы развернуть траекторию цифровизации к технооптимистическому сценарию.

Литература

Батищев, Г. С. (1963). *Противоречие как категория диалектической логики*. Москва. Высшая школа. 120 с.

Кизима, В. В. (2010). Постнеклассические практики: рефлексивность и управление. *Вопросы философии*. 2010. № 3. С. 54–65.

Rapaport, W. J., Koenig, D. R. (1995). Ontology of Cyberspace: Questions and Comments. *ResearchGate* [Электронный ресурс]. URL:

https://www.researchgate.net/publication/2458574_THE_ONTOLOGY_OF_CYBERSPACE_Questions_and_Comments (дата обращения: 31.05.2025).

«САМОРАСШИРЕНИЕ» ЧЕЛОВЕКА
В КОНТЕКСТЕ ТРАНСФОРМАЦИЙ
АНТРОПОСФЕРЫ БУДУЩЕГО

Т. Г. Лешкевич

Южный федеральный университет

Leshkevicht@mail.ru

Анализ проблематики антропологии будущего нацеливает на учет как новых возможностей, предлагаемых искусственным интеллектом (ИИ), свидетельствующих о «саморасширении» человека, так и негативов технологических новаций, которые фиксируют амбивалентность конфигураций будущего. Тем не менее, английское слово *enhancement*, используемое для описания субъектной перспективы техносимбиоза, переводится как улучшение и позволяет вести речь о том, что технореальность будущего предполагает трансформации человека на основе его «сращенности» и гипервзаимосвязанности с ИИ. Одобрение техносимбиоза выливается в формулу, предложенную Э. Кларком: чтобы «познать себя, знай свои технологии» [Clark, 2004, p.184]. Положительное отношение людей к результатам практик улучшения человека, обусловленных тождественностью эффектов технологического и нейробиологического функционирования, известно всем. Имеются в виду «импланты: от электронных кардиостимуляторов, имплантов суставов, приборов, улучшающих слух, зрение и до чипов, внедряемых в мозг и расширяющих функциональность, в том числе и так называемых корковых имплантов – «усилителей» ощущений, или

же аудио имплантов в зуб, разработанных в 2002г. и передающих звуковую информацию из зуба во внутреннее ухо по костному трансиверу» [Лешкевич, 2025, с. 44]. Есть разработки улучшенных биоинспирированных материалов с учетом принципов живой природы для костных имплантов. Подобные аддитивные биогибридные инновации выступают приоритетами будущего. При этом значение имеют как «сильные» формы усовершенствования с долгосрочными либо постоянными результатами, так и формы «временные», обеспечивающие краткосрочные изменения [Гребенщикова, 2016, с. 30]. Сильные свидетельствуют о тенденции развития сложных гибридов, объединяющих человеческое и нечеловеческое. Слабые формы, высвечивают иной эффект, и, в частности то, что триумф технологий, направленных на вторжение в мир живого, «срашенность» с технологиями, может сопровождаться негативными побочными последствиями. В этом контексте вопрос Дж. Нейсбита: «Спасет технология человечество или уничтожит его?» – не теряет своей актуальности [Нейсбит, 2005].

В отношении когнитивных процессов следует отметить, что активное замещение индивидуальных мыслительных поисков алгоритмизированными решениями, свидетельствует о переустройстве процесса «производства» мысли. Вытеснение понятийного мышления «мышлением технологиями» с доминированием обязательного набора цифровых функций, превращающим человека в актанта, находится в конфронтации с выводами когнитологов о том, что именно понятийное мышление выступает центральным фактором в процессе формирования человеческого интеллекта [Холодная, 2020, с. 21]. Тем более, что «спайка» человеческого интеллекта с возможностями ИИ,

партнерство с Chat GPT, ставшее уже рядовой процедурой, порождают эффект «ленивого мозга». Привычка к заимствованию из Сети готовых решений приводит к эрозии самостоятельного мыслительного аппарата и чревата утратой человеческой самостоятельности в целом. Более того стремительно развивающийся Chat GPT и ИНС обнаруживают беспомощность социальных амортизаторов, которые в состоянии противостоять усилению их диктатуры. Нет и убедительной аргументации относительно того, как социогуманитарная рефлексия может уменьшить риски произвольного и бесконтрольного функционирования системы «нейроинтерфейс мозг-компьютер». Тем более, что современная нейросеть демонстрирует субъектоподобные качества и протоцелесообразность, в отношении которой отмечен тот факт, что ИНС могут «ставить и менять цели исходя из возможности, самостоятельно трансформировать собственную техническую и программную «архитектуру», учитывая скрытые свойства компонентов, софтов и алгоритмов» [Игнатъев, 2021, с. 146].

Современный методолог Дж. Ло обосновывает значимость сопоставления двух полюсов: «здесь-внутри» и «там-вовне», которое имеет вес для антропосферы будущего. В области «здесь-внутри» мы выступаем как субъекты со всеми психофизиологическими особенностями, а «там-вовне» мы предстаем «как сети значимых и материальных отношений [Ло, 2015, с. 144]. Так проинтерпретированная связка между «здесь-внутри» и «там-вовне» высвечивает канал, указывающий на отношения «вход-выход» с требованием, состоящем в том, чтобы технологический артефакт был адаптивен органике. Добавим, что в пространстве «здесь-внутри» располагается внутренняя

«лаборатория» когнитивных процессов, кодируются нейронные сигналы, определяется сеть семантических ориентаций, дает о себе знать сапиентность – разумность человека.

Расширение человеческих функциональных возможностей и супертехнологическое увеличение ресурсов организма часто оценивается как своеобразный выход за пределы «нормального» и вызывает вопрос: предпочтут ли люди естественным органам артефакты и возрастающую зависимость от машин? Озабоченность вызывает и то, что переход к цифровой «грамматике бытия» чреват утверждением в качестве ведущего паттерна существования компьютерно-опосредованное поведение и онлайную проекцию разнообразных сфер жизни. А это ведет к метарепрезентативности телесности, некоему бестелесному опыту, поставляемому ИИ и ИНС. Эта ситуация, обеспеченная запрограммированными канальными опциями, устрашает перспективой потери человеческих навыков тактильной связи друг с другом. И если традиционные «мы-отношения», как правило, осуществлялись «глаза в глаза», то посредничество Интернета сделало доминирующей цифровую коммуникацию, девальвирующую значимость непосредственного взаимодействия и физических контактов. Более того, метателесность может порождать различные киборгизированные формы, превосходящие биологические ограничения и связанные с выходом «за пределы возможностей естественного существования человека к гибриднему типу функционирования, дополняющему нейрофизиологические функции технологическими возможностями» [Лешкевич, 2022, с. 55].

Таким образом, в контексте проблематики антропосферы будущего на фоне обсуждения проектов искусственной личности, актантов и феноменов non-humans фундаментом становятся самоорганизующиеся интеллектуальные устройства на небиологических носителях, демонстрирующих квазисубъектность, сверхисполнительность и ассоциативность. Делается вывод, что «саморасширение» на пути техносимбиоза свидетельствует о новом выборе современного человека, достраивающего себя посредством технологий. Но если мыслить о будущем с целью сохранения человекоразмерного бытия, то необходима система мер, среди которых: (1) установка на маркировку контента, сгенерированного Chat GPT; (2) самоограничение на использование ИИ-продукции, подавляющей высшие функции человеческого мышления; (3) публичные дискуссии, вскрывающие риски и угрозы поглощенностью ИИ-взаимодействиями.

Литература

Гребенщикова, Е. Г. (2016). Проекты улучшения человека и тезис технологической неизбежности. *Человек*. № 5. С. 30-39.

Игнатъев, В. И. (2021). Проблема техносубъекта: о субъектности «сущностей-конструкторов». *Идеи и идеалы*. Т. 13. № 1. Ч. 1. С. 130–150.

Лешкевич, Т. Г. (2022). Человек-виртуал и передача культурных ценностей поколению эпохи цифры. *Вопросы философии*. № 3. С. 53–63.

Лешкевич, Т. Г. (2025). Субъектоподобные качества ИИ: «стыковка» humans и non-humans. *Вопросы философии*. № 4. С. 39–47.

Ло, Дж. (2015). *После метода. Беспорядок и социальная наука*. М.: Изд-во Институт Гайдара.

Нейсбит, Дж. (2005). *Высокая технология, глубокая гуманность: Технологии и наши поиски смысла*. М. АСТ; Траюиткнига,

Холодная, М. А. (2020). Многомерная природа показателей интеллекта и креативности: методические и теоретические следствия. *Психологический журнал*. Том 41. № 3. С. 18–31.

Clark, A. (2004). *Natural-Born Cyborgs: Minds, Technologies, and the Future of Human Intelligence*. Oxford, England: Oxford University Press.

О ПРИРОДНОЙ И НОРМАТИВНОЙ РЕАЛЬНОСТИ В НЕОРАЦИОНАЛИЗМЕ Р. БРАССЬЕ

О. А. Лунев-Коробский

ИФПР СО РАН, г. Новосибирск

sickrrett@gmail.com

*Статья подготовлена при финансовой поддержке РФФ,
проект № 25-28-01064 «Трансформация рациональности:
от инструментального разума к неорационализму».*

Спекулятивный поворот – это течение в современной философии. Минимальный критерий для отнесения философского проекта к спекулятивному повороту – стремление к преодолению антропоцентризма как предпосылки теоретизирования. Неорационализм – одно из ответвлений этого многогранного и продолжающего развитие комплекса философских концепций¹. Они связаны единой квазиметодологической предпосылкой – реабилитировать спекулятивную метафизику как теоретический дискурс о мире, который после «коперниканской революции» Канта считался недоступной для человека сферой ноуменов.

Отсюда главный предмет интереса философии спекулятивного поворота, который Мейясу назвал проблемой

¹ Об этом свидетельствуют академические и параакадемические ресурсы: журналы Collapse и Angelaki, издательские программы Edinburgh University Press, re.press, Schism press, Urbanomic, интернет-проекты deontologistics.co, tophilosophy.com, uberty.org и др.

«корреляционизма». Как получить доступ к миру самому по себе, при том условии, что мы всегда вынуждены искать к нему доступ из позиции мира для нас [Мейясу, 2015, с. 12]¹.

В философии спекулятивного поворота выделяются две основные стратегии проработки проблемы доступа к реальности: стратегия *концептуального творчества* и стратегия *рациональной критики*². Далее я реконструирую на конкретном примере одну из них – стратегию рациональной критики, которая характерна для неорационализма в целом и для Р. Брассье как одного из его ведущих представителей в частности³.

В статье “The view from nowhere” [Brassier, 2011] Брассье синтезирует концепцию синоптической философии У. Селларса

¹ «<...> корреляционизм сам не занимает никакую позицию, будь то иррациональную, религиозную или поэтическую: он не производит положительный дискурс об абсолюте, он ограничивается осмыслением пределов мышления, которые являются границей языка, постижимой лишь с одной стороны. Корреляционизм не предлагает положительного обоснования определённой религиозной веры, но он подрывает любую попытку рационально лишить веру легитимности на основании того, что её содержание немыслимо» [Мейясу, 2015, с. 56].

² В духе Делёза и Канта соответственно. Ср.: «<...> философия – это искусство формировать, изобретать, изготавливать концепты» [Делёз, Гваттари, 2009, с. 6]. Также: «Только трезвая и строгая, но справедливая критика может освободить от этой догматической иллюзии, привлекающей воображаемым блаженством столь многих людей к теориям и системам, и ограничить все наши спекулятивные притязания только сферой возможного опыта <...> с помощью произведённого согласно достоверным основоположениям определения границ его [разума], с величайшей надёжностью надписывающего своё *nihil ulterius* на поставленных самой природой геркулесовых столбах <...>» [Кант, 2006б, с. 493–495]. Однако: «Раз с помощью критики нашего разума мы в конечном счёте узнаём, что при чистом и спекулятивном употреблении разума мы, собственно, ничего не можем познать, то не должна ли она поэтому открыть более широкое поприще для *гипотез* : ведь если нельзя что-то утверждать, то, может, по крайней мере, будет позволительным [что-то] придумывать и [о чём-то] высказывать мнение?» [Кант, 2006а, с. 971].

Эти стратегии не являются взаимоисключающими, одни представители спекулятивного поворота делают ставку скорее на концептуальное творчество, чем на рациональную критику, и наоборот.

³ Реконструкция стратегии концептуального творчества, а также сравнительный анализ двух этих стратегий выходят за рамки данной работы. Исследование будет продолжено.

с немоцентризмом Т. Метцингера, чтобы подвергнуть критике современные концепции самости как проявления мифа о данном.

Брассье утверждает, что теория коммуникативной рациональности Ю. Хабермаса ограничена резким размежеванием сфер нормативного и природного. Согласно Хабермасу, опредмечивание сознания концептуально невозможно, поскольку объективация игнорирует сущность сознания, которая становится возможной только в социальной коммуникации.

По мнению Брассье, Хабермас воспроизводит картезианский дуализм субстанций, потому что не проводит адекватное концептуальное различие между нормативно обусловленным индивидом и феноменологической реальностью субъекта. Это ведёт к неспособности Хабермаса «артикулировать различие между теоретическим опредмечиванием и дискурсивной практикой» [Brassier, 2011, p. 19].

Согласно Брассье, путь к «всестороннему пониманию мира и нашего места в нём» лежит через концептуальную интеграцию предварительно разграниченных измерений нормативного и природного, потому что ключевой компонент рациональности – это субъект-носитель рациональности как физический организм, реализующий в практике (в том числе коммуникативной по Хабермасу) сферу теоретических трансформаций, которую этот же субъект динамически концептуализирует.

Другими словами, рациональная самость проявляет себя во взаимодействии со средой, но в другом качестве та же рациональная самость есть эмпирическая субъективность, а наш лучший инструмент для понимания всего, что дано нам в опыте, – это натурализм. Поэтому понимание рациональности

невозможно без сочетания анализа оформляемой концептуальными нормами практики и объяснения природы в соответствии с каузальными закономерностями физического мира. Таким образом, самость – это не сущность, а *locus* рационального и физического субъекта действия¹.

Далее Брассье обращается к феноменологической традиции понимания субъекта. Он утверждает, что гипостазирование интенциональности приводит к надуманному эпистемологическому препятствию между самостью как трансцендентальным резервуаром опыта и миром за пределами этого опыта.

Феноменология – это доктрина описания опыта из позиции предполагаемой чистой самости как формы, в которой дан всякий опыт. Тем самым феноменолог сводит субъективность агента до уровня эмпирической данности как содержания самости: самость трансцендентально первична и обуславливает всякий опыт, она структурно необходима, как неизменная форма опыта.

Суть критики Брассье в адрес позиции феноменологии по вопросу о субъекте такова: феноменолог «не учитывает различие между интенциональными феноменами и феноменами сознания <...> не приводит никаких аргументов в пользу заявления, что феноменологическая “данность” остаётся несводимой к психологическому и/или когнитивному опыту» [Brassier, 2011, p.

¹ Причём, вслед за Селларсом, Брассье понимает действие в широком смысле: «"истина" – это знак *того, что нечто подлежит исполнению* [that something is to be done], поскольку умозаключение является действием [doing]» [Brassier, 2011, p. 7]. Ср.: «<...> критика легко обнаружит догматическую видимость и заставит чистый разум отказаться от чрезмерных притязаний в спекулятивном применении и возвратиться в пределы своих владений, а именно в область практических принципов» [Кант, 2006а, с. 999].

21]. Феноменолог допускает концептуальную небрежность, почти идентичную «заблуждению философа» Хабермаса: в первом случае – феноменологическая интенциональность есть *сущность*, делающая легитимным всякий опыт, во втором – коммуникативность есть *первостепенная, сущностная черта* человеческой рациональности.

Немоцентризм Метцингера становится для Брассье отправной точкой развенчания феноменологической позиции. В работе 2004 г. “Being No One: The Self-Model Theory of Subjectivity” Метцингер предлагает детализированную концепцию, ядро которой – тезис, что самость представляет собой специфический аспект опыта.

Данность опыта от первого лица – это конструктивная издержка нашего мозга. Чтобы избежать непомерных энергетических затрат на моделирование среды, сознание устроено так, что представляет собой нечто вроде феноменального экрана для проецирования данных извне. Мозг человека структурно не способен заглянуть по ту сторону этого экрана. Он от природы устроен так, чтобы решать биологические задачи наиболее экономным и эффективным способом, то есть транслировать и обрабатывать необходимую для функций самосознания и выживания в целом информацию, и при этом избегать бесконечной рекурсии, которая неизбежна при стремлении смоделировать сам процесс моделирования феноменального опыта. Феноменальная самость – это артефакт эволюции, а наш опыт самости – «тень» на стене «пещеры, в которой никого нет» [Brassier, 2011, p. 16].

Достоинство немоцентрической концепции Метцингера заключается в том, что её концептуальные элементы согласуются

с нейробиологическими и психофизиологическими данными экспериментальной науки. Разумеется, немоцентризм остаётся гипотезой, однако она предпочтительна в силу того, что натурализует, и тем самым редуцирует (*explains out*) самость. Вопреки опасениям Хабермаса, это не означает, что человек оказывается бессодержательным концептом. Брассье предлагает избавиться от мифа о том, что самость дана нам как некая *сущность*. Самость – это режим функционирования опыта сознания, который потенциально *объясним с помощью науки*¹.

Литература

Делёз, Ж., Гваттари, Ф. (2009). *Что такое философия?* Москва. Академический Проект.

Кант, И. (2006а). *Сочинения на немецком и русском языках. Т. 2: Критика чистого разума: в 2 ч. Ч. 1.* Москва. Наука.

Кант, И. (2006б). *Сочинения на немецком и русском языках. Т. 2: Критика чистого разума: в 2 ч. Ч. 2.* Москва. Наука.

Мейясу, К. (2015). *После конечности: Эссе о необходимости контингентности.* Екатеринбург; Москва. Кабинетный учёный.

¹ В сущности, аргументация Брассье сводима к хорошо известному *lex parsimoniae*: если есть гипотеза, которая предполагает более простое объяснение, то в большинстве случаев она предпочтительнее гипотезы, которая предполагает более сложное объяснение. Такой редукционизм, внимательный к концептуальным различиям, является характерной чертой работ неорационалистов, к которым вместе с Брассье причисляют П. Вульфендейла и Р. Негарестани. См., напр., Negarestani, R. (2018). *Intelligence and Spirit*. Urbanomic; Wolfendale, P. (2014). Is there a TV in my head?: Content, Functional Mapping, and the Myth of the Given. *Taala: Quaderns de pensament*. No. 46. P. 27–38; Wolfendale, P. (2019). The Reformatting of Homo Sapiens. *Angelaki*. Vol. 24. No. 1. P. 55–66.

Brassier, R. (2011). The View from Nowhere. *Identities. Journal for Politics, Gender and Culture*. Vol. 8. No. 2. P. 7–23.

Цифровое посмертие: философские и технологические аспекты «загробной жизни» в цифровую эпоху

Е. Г. Миляева

Челябинский государственный университет, Челябинск
morning_demon@mail.ru

Концепция цифрового посмертия – сохранения человеческого сознания, воспоминаний или идентичности в цифровой форме после биологической смерти – бросает вызов традиционным представлениям о смертности и личности. Укорененная в философии трансгуманизма и ставшая возможной благодаря развитию искусственного интеллекта (ИИ), эта идея ставит сложные вопросы о том, что значит существовать, помнить и взаимодействовать в «постбиологическом» мире.

Цифровое посмертие тесно связано с трансгуманизмом – движением, выступающим за использование технологий для преодоления биологических ограничений человека, включая смерть [Savin-Baden, 2020; Arasu, 2023]. В рамках трансгуманизма цифровое бессмертие рассматривается как логичный этап эволюции, при котором сознание переносится на небιологические носители. Это согласуется с такими проектами, как «Россия 2045», ставившим целью достижение «кибернетического бессмертия» через клонирование сознания и ИИ [Savin-Baden, 2020]. С философской точки зрения это переопределяет смерть не как конец, а как переход в цифровое

существование, стирая границы между жизнью, памятью и симуляцией [Yasseri, 2022; Park, 2024].

Исследовав представленные в Сети сайты (HereAfter AI, Replika и др.), предлагающие услуги посмертия, можно уверенно сказать, что сегодня цифровые аватары (часто называемые «боты-покойники» или «боты скорби») – всего лишь симуляции, лишенные субъективного сознания своих биологических прототипов. Этические вопросы, связанные с созданием посмертных ИИ-агентов (например, риск эксплуатации или психологического вреда для скорбящих), подчеркивают необходимость регулирования, которое уравновесит технологические амбиции и уважение к человеческому достоинству [Hu, 2024].

Современные технологии цифрового посмертия используют генеративный ИИ, воссоздающий личность на основе цифрового следа – постов в соцсетях, электронных писем и мультимедиа [Hu, 2024]. Стартапы вроде HereAfter AI и проекты типа Replika применяют большие языковые модели (LLM) для создания интерактивных аватаров, имитирующих речь и знания умерших [Biano, 2024]. Например, проект «Dadbot» сохранил голос и воспоминания отца разработчика, используя сотни часов записей его разговоров, что позволило продолжать общение после его смерти.

Технологии Building Information Modeling (BIM) и виртуальная реальность (VR) расширяют возможности, позволяя оцифровывать физические пространства и культурное наследие [Huberman, 2018; Elkhova, 2024]. Хотя они делают мемориализацию доступнее, существует риск коммодификации

горя, особенно на платформах, предлагающих «услуги цифрового посмертия» по подписке [Hu, 2024; Narju, 2024].

Несомненно, что коммерческая индустрия цифрового посмертия столкнется с вопросом: «Должны ли люди сохранять контроль над своими цифровыми «я» после смерти?» В текущей практике часто отсутствуют четкие механизмы получения согласия как от умерших («доноров данных»), так и от живых, взаимодействующих с их цифровыми аватарами. Для одних ИИ-аватары становятся источником утешения, но у других могут вызывать затяжную скорбь или экзистенциальный дискомфорт – особенно если возникает сомнение: общаются ли они с подлинной личностью или всего лишь с алгоритмической иллюзией. Среди предлагаемых решений – ограничение подобных сервисов для взрослых, обеспечение прозрачности в использовании данных для обучения ИИ, а также разработка процедур «завершения» работы цифровых аватаров. Например, обсуждаются концепции «чувствительного вывода из эксплуатации» ботов-покойников – по аналогии с цифровыми похоронами.

Феномен цифрового посмертия ставит перед цифровой философией и антропологией будущего радикально новые вопросы, требующие пересмотра самих оснований человеческого бытия. Технологическая возможность сохранения и воспроизведения сознания разрушает традиционные представления о личности, памяти и смерти, заставляя философию и антропологию искать новые концептуальные рамки для осмысления «постбиологического» существования. Серьезнейший вызов заключается в необходимости переосмыслить саму природу человеческого. Где проходит та

неуловимая грань между живой личностью и ее цифровой реконструкцией? Как определить подлинность сознания, когда алгоритмы могут воспроизводить наши мысли, эмоции и манеру общения? Эти вопросы требуют принципиально новых эпистемологических подходов. Не менее сложной представляется этическая проблематика цифрового бессмертия. Технологический прогресс опережает нашу способность осмыслить его последствия для человеческого достоинства, скорби и отношений между поколениями. Социальные последствия также вызывают серьезные опасения – каким образом изменится структура общества, когда ушедшие поколения останутся в цифровом пространстве как активные собеседники?

Цифровая философия и антропология будущего стоят перед необходимостью выработать новые методологические инструменты для анализа этих трансформаций. Сегодня человечество имеет дело не просто с очередным технологическим скачком, а с фундаментальным антропологическим переворотом, где сама природа человеческой субъективности становится объектом цифрового моделирования.

Главная философская задача ближайших десятилетий – сохранить гуманистическое измерение в условиях, когда технологии обещают преодолеть биологическую конечность человека. Речь идет не только о технических возможностях, но о необходимости переосмыслить, что делает нас людьми в эпоху, когда сознание потенциально может существовать вне органического тела. Этот вызов требует глубокого междисциплинарного синтеза, объединяющего философию

рефлексию, антропологический анализ и этическое регулирование технологического развития.

Литература

Arasu, P. (2023). Lovecraftian Games: The Afterlife of Cthulhu on Valve's Steam Client. *Games and Culture*, 0(0). n. pag.

Biano, I. W. (2024). The digital afterlife between hypersecularization and postsecularity in contemporary speculative TV seriality: Devs, Upload and Black Mirror's 'San Junipero'. *Australasian Journal of Popular Culture*. n. pag.

Elkhova, O. I. (2024). Eternal codes: philosophical reflections on "cybernetic immortality". *Vestnik of the Kyrgyz-Russian Slavic University*. n. pag.

Harju, A. A. (2024). Theorising Digital Afterlife as Techno-Affective Assemblage: On Relationality, Materiality, and the Affective Potential of Data. *Social Sciences*. n. pag.

Hu, Y., Jin, Z., & Chen, Q. (2024). The digital preservation of national memorial architecture: A new approach with BIM and VR technology. *The Art & Design Research*. n. pag.

Hollanek, T., & Nowaczyk-Basińska, K. (2024). Griefbots, Deadbots, Postmortem Avatars: on Responsible Applications of Generative AI in the Digital Afterlife Industry. *Philosophy & Technology*. n. pag.

Huberman, J. (2018). Immortality transformed: mind cloning, transhumanism and the quest for digital immortality. *Mortality*. No. 23. P. 50–64.

Park, L. Y. J., Oh, Y. W., & Sang, Y. (2024). Digital Access, Digital Literacy, and Afterlife Preparedness: Societal Contexts of Digital Afterlife Traces. *Social Media + Society*. 10(3). n. pag.

Savin-Baden, M., & Mason-Robbie, V. (Eds.). (2020). *Digital Afterlife: Death Matters in a Digital Age*. Chapman and Hall/CRC. n. pag.

Yasseri, T., Gildersleve, P., and David, L. (2022). *Collective Memory in the Digital Age*. Collective Memory, Elsevier. n. pag.

СОВРЕМЕННОЕ ИСКУССТВО И НОВЫЙ АНТРОПОГЕНЕЗ: ОТ НОМО SAPIENS К НОМО DIGITALIS?

С. А. Новинская

Уральский федеральный университет имени Б.Н. Ельцина,

Екатеринбург

[@gmail.com](mailto:novinson06@gmail.com)@ygmail.com

В последние десятилетия понятие сознания становится все более многогранным и сложным для однозначного определения. Традиционно сознание рассматривалось как совокупность субъективных переживаний, осознания себя и окружающего мира, а также способности к мышлению и рефлексии. Однако с развитием технологий и проникновением цифровых средств в повседневную жизнь границы между внутренним миром человека и внешней средой начинают стираться. Мы живем в момент фундаментального сдвига. Технологии больше не просто расширяют наши возможности – они меняют саму основу человеческого опыта. Интернет, дополненная реальность, цифровые аватары, искусственный интеллект – все это формирует новый тип субъекта. На смену *Homo sapiens*, «человеку разумному», приходит *Homo digitalis* – человек, существующий в цифровой среде, взаимодействующий с миром через интерфейсы и алгоритмы.

Все большая часть человеческой активности, коммуникаций и даже идентичности переносится в виртуальные пространства, что существенно трансформирует наше

восприятие себя и окружающей реальности. Одновременно с этим меняется и наше телесное существование: виртуальная реальность формирует новые формы взаимодействия тела и сознания, расширяя традиционные рамки человеческого опыта. Антропогенез в эпоху информационных технологий формирует новую ценностную парадигму и влияет на культурное развитие человека, трансформируя его биологические и социальные характеристики.

Вопрос не только в том, как технологии меняют повседневную жизнь, но и в том, как они трансформируют само понимание человека. И если философы размышляют об этом концептуально, то современное искусство делает эти изменения видимыми и чувственно переживаемыми. Оно становится зеркалом, в котором можно разглядеть очертания новой антропологической формы.

В связи с этим особенно актуальным становится то, как современное современное искусство отражает и осмысляет эти процессы перехода человека в виртуальный мир. Художники, работающие с цифровыми технологиями, создают не просто визуальные образы, но и концептуальные пространства, в которых исследуются новые формы сознания, идентичности и телесности. В данной статье мы рассмотрим, каким образом современное искусство интерпретирует феномен виртуализации человеческого опыта и какие перспективы развития этой тенденции видят современные творцы.

Концепт Homo digitalis начал формироваться в контексте кибернетических, сетевых и постгуманистических теорий. Французский философ Пьер Леви описывал цифрового человека как существо, обитающее в «виртуальной коллективности»,

основанной на обмене знаниями и символами. Позднее Борис Гройс и Юваль Харари указывали на то, что цифровая среда кардинально меняет структуру субъекта: теперь это не автономное, замкнутое «я», а расплывчатая, фрагментированная идентичность, представленная в виде аватаров, профилей и потоков данных.

В своей книге «Homo Digitalis: Smartphones, soziale Netzwerke und das Gehirn» Кристиан Монтаг определяет Homo digitalis как «человека, который существует одновременно в физическом и цифровом пространстве». Его тело может быть незаметным или вовсе отсутствовать (в виртуальной среде), но его цифровое присутствие постоянно актуализируется. Он мыслит не последовательно, а через гиперссылки и интерфейсы. Его память все чаще передана внешним носителям: от облачных хранилищ до нейросетей.

Творчество китайской художницы Лу Ян (Lu Yang) представляет собой уникальный синтез буддийской философии, нейронауки и цифровых технологий, что позволяет ей исследовать переход человеческого сознания в виртуальную реальность через призму трансформации идентичности, кибернетического бессмертия и множественности «Я». Ее проекты, такие как DOKU the Self и Electromagnetic Brainology, служат визуальными манифестациями этих процессов.

Например, в своей серии DOKU («Digital Reincarnation») Лу Ян переосмысливает буддийскую концепцию сансары через призму цифровых технологий. Проекты вроде DOKU – Digital Alaya отсылают к алая-виджняне – «сознанию-хранилищу» в буддийской философии, которое художница трансформирует в цифровой архив данных. Здесь реинкарнация представлена как

бесконечный цикл загрузки/выгрузки сознания в виртуальные миры, где аватар становится новой формой существования.

В проекте *Electromagnetic Brainology* Лу Ян исследует взаимодействие мозга и машин. Электроэнцефалография (EEG) и интерфейсы «мозг-компьютер» визуализируют нейронную активность как психоделические паттерны, переводя биологические процессы в цифровые артефакты. Работа *Brain Control Messenger* демонстрирует, как технологии могут манипулировать эмоциями и восприятием, создавая искусственные состояния сознания.

Другой проект «*Pulse Room*» художника Рафаэля Лозано-Хеммера, созданный в 2006 году, представляет собой интерактивную инсталляцию, где физиологический ритм человека становится основой для коллективного цифрового архива. Хотя работа существует в физическом пространстве, ее концепция метафорически отражает переход человеческого сознания в виртуальную реальность через кибернетизацию телесности и трансляцию биоданных в эстетический объект. Участники, прикасаясь к сенсору, передают свой пульс в систему, которая преобразует его в световые импульсы ламп. Это создает «цифровой отпечаток» биологического процесса, визуализируя сердцебиение как данные, которые сохраняются в инсталляции. Каждая лампа становится носителем чьей-то временной «жизни» – подобно тому, как аватары в VR хранят информацию о пользователях. Проект реализует идею обратной связи между телом и технологией, где биологический сигнал (пульс) становится управляющим параметром для арт-объекта. Это отражает процесс гибридизации человека и машины – ключевой аспект перехода к *Homo digitalis*.

Проект Сандры Перри «Graft and Ash for a Three Monitor Workstation» – это глубокая рефлексия над тем, как переход в виртуальную реальность меняет представления о теле, идентичности и сознании. Физическое усилие зрителя становится «топливом» для цифрового пространства, подчеркивая, что даже в виртуальной реальности тело остается активным участником процесса. Цифровой аватар Перри на экранах – не идеальная копия, а намеренно искаженное, фрагментированное отражение ее личности. Аватар рассказывает о невозможности полностью воспроизвести тело, цвет кожи, черты лица в цифровой среде, что особенно важно для художницы как чернокожей женщины. Это подчеркивает ограниченность технологий в передаче уникального человеческого опыта и иллюстрирует, как в виртуальной реальности идентичность становится множественной и подверженной постоянной трансформации. Таким образом, Перри демонстрирует, что виртуальное пространство не отменяет, а трансформирует социальные, расовые и телесные различия, делая их частью нового цифрового опыта.

Таким образом, современное искусство выступает не только зеркалом эпохи, но и лабораторией, где моделируются сценарии будущего антропогенеза. Через художественные практики становится очевидно, что *Homo digitalis* – это не просто человек, освоивший новые технологии, а субъект, существующий на стыке биологического и цифрового, реального и виртуального, индивидуального и коллективного. Искусство выявляет как возможности, так и риски этого перехода: расширение границ сознания и тела, появление новых форм идентичности, но также и новые вызовы, связанные с утратой

автономии, фрагментацией «я», усилением социального контроля и неравенства.

В итоге, анализ художественных стратегий и проектов, посвященных виртуализации человеческого опыта, позволяет сделать вывод о том, что современное искусство не только фиксирует момент перехода к Homo digitalis, но и формирует критическую оптику, необходимую для осознания и гуманизации этого процесса. Оно помогает увидеть, что в цифровую эпоху вопросы тела, сознания, памяти и идентичности становятся не только технологическими, но и глубоко гуманитарными, требующими нового философского и этического осмысления.

Источники

Андреева, А. (2021). Цифровое бессмертие: есть ли жизнь после смерти в компьютере. *РБК.ru*. [Электронный ресурс]. URL: <https://trends.rbc.ru/trends/futurology/5e2825d29a7947f9fe755f87> (дата обращения: 10.03.2025)

Lozano-Hemmer, R. Pulse Room. *ATELIER LOZANO-HEMMER* [Электронный ресурс]. URL: https://www.lozano-hemmer.com/pulse_room.php (дата обращения: 15.03.2006).

Perry, S. (2016). Graft and Ash for a Three Monitor Workstation [Электронный ресурс]. URL: <https://sondraperry.com/Graft-and-Ash-for-a-Three-Monitor-Workstation> (дата обращения: 27. 03. 2025.)

Yang, L. (2023). DOKU the Self [Электронный ресурс]. URL: <http://luyang.asia/2023/05/20/doku-the-self/> (дата обращения: 23.04.2025).

Yang, L. (2017). Electromagnetic Brainology [Электронный ресурс]. URL: <https://radii.co/article/watch-lu-yangs-electromagnetic-brainology-brain-control-messenger> (дата обращения: 23.04.2025).

Литература

Белл, Д. (2004). Грядущее постиндустриальное общество. Опыт социального прогнозирования. М. Academia.

Горюнов, И. А. (2018). Постмодерн: от интернета вещей к нейронету людей? *Философия хозяйства*. № 3 (117). С. 92–103.

Гройс, Б. (2023). Политика поэтики. М. AdMarginem.

Ермилов, К. А. (2021). Постчеловек, искусственное тело и пределы трансгуманизма. *Acta Eruditorum*. № 37. С. 47–50.

Тоффлер, Э. (2010). Третья волна. М. АСТ.

Харари, Ю. В. (2018). *Homo Deus: Краткая история завтрашнего дня*. М. Синдбад.

Lévy, P. (1987). *La Machine Univers. Création, cognition et culture informatique*. Paris. La Découverte.

Lévy, P. (1995). *Qu'est-ce que le virtuel?* Paris. La Découverte.

Lévy, P. (1997). *Cyberculture. Rapport au Conseil de l'Europe dans le cadre du projet «Nouvelles technologie: coopération culturelle et communication»*. Paris.

Montag, C. (2018). *Homo Digitalis: Smartphones, soziale Netzwerke und das Gehirn*. Berlin. Springer.

ВИРТУАЛЬНОСТЬ И РЕАЛЬНОСТЬ ИНТЕРНЕТА ВЕЩЕЙ: ТЕХНИЧЕСКИЙ, ПСИХОЛОГИЧЕСКИЙ И СОЦИАЛЬНО- ФИЛОСОФСКИЙ АСПЕКТЫ

Е. В. Ускова

Уральский федеральный университет имени Первого президента
России Б.Н. Ельцина, Екатеринбург
e.v.uskova80@mail.ru

Термин «интернет вещей» (IoT) был предложен Кевином Эштоном. Этим термином он обозначил глобальную электронную сеть, соединяющую физические объекты с Интернетом. ([Ashton, 2009]) «IoT есть глобальная кибер-физическая система, которая состоит из различных физических объектов, подключенных к глобальной информационной среде посредством искусственных подсистем, соответствующего программного обеспечения, электронных платформ, управляющих контроллеров и коммуникационных адаптеров». [(Курбатов, Папа, 2023, с. 49)] Сам IoT может пониматься как совокупность технических артефактов – это датчики, сети (коммуникация), аналитика (облака), приложения. Либо может пониматься и как социальный интернет вещей (SIoT), в этом случае мы имеем в виду не только техническую возможность связи разных устройств в единую сеть, но включаем в это взаимодействие человека, поскольку именно он является причиной, целью и тем самым «пользователем», ради которого выстраиваются данные сети. Устройства, используемые людьми

в самых разных областях, стали получать обозначение «умные», отсюда появляются «умные дома», «умные города», «умные фабрики» и т.д. По некоторым данным к единой сети Интернет в 2025 году должно быть подключено до 75 млрд вещей. Все эти вещи генерируют огромное количество информации, которая нуждается в постоянно обработке и использовании, что приводит к появлению отдельной отрасли знания, и профессии, связанной с ней – Big Data.

Если мы посмотрим на эту новую техническую реальность с психологической точки зрения, то необходимо будет понять, какую роль играют новые помощники человека в его жизни и отношениях с другими людьми, как их меняют и перестраивают. Для того, чтобы у пользователей была возможность взаимодействия с «умными» устройствами, выполняется установка соответствующих приложений на смартфон или любой другой гаджет, а также возможно управление через голосовых помощников. Соответственно, при использовании «умных» вещей, мы можем осуществлять управление через нажатие кнопок или голосом. Если мы общаемся с нашей «умной» колонкой, чайником, холодильником, выключателем света или любым другим датчиком, то невольно переносим на него все стандартные атрибуты собеседника, что ставит перед нами вопрос о статусе «умных» устройств: кто они? Объекты или субъекты? А от ответа на вопрос зависит и наш способ общения с этими устройствами, а также возможность их включение в рутину повседневности. Люди часто не хотят впускать эти устройства в свою жизнь, потому что не желают менять свои привычки (говорить вместо того, чтобы нажимать кнопку или выключатель). Очевидно, что с этой проблемой сталкиваются

люди старшего поколения, молодые воспринимают это как данность.

То есть с психологической точки зрения значимость новых технологий, которые приходят в нашу жизнь, проистекает не только, а может быть не столько из тех функций, которые заложили разработчики в данную вещь, а из того, как мы сможем приспособить эти функции для себя или найти этой «умной» вещи то применение, которое необходимо именно нам. Поэтому, на первый план часто выходит не польза вещи, как таковая, «...но и представление о том, как к этой технологии относятся значимые другие, и соответствует ли она сложившимся социальным нормам, а также оценку самоэффективности, т.е. собственной способности контролировать достижение результата, например, оценку достаточности имеющихся знаний и навыков для работы с новым устройством». [(Журавлёва, Нестик, 2019, с. 39)]

Философский аспект интернета вещей. Происходит изменение социальной реальности под воздействием IoT, что выражается в трансформации её границ и свойств (расширение и углубление). Люди рождаются, вырастают и социализируются в совершенно другой среде, нежели ранее; меняются их представления о своём теле и его границах; сознании и внутреннем мире; реальности и виртуальности. Вещи, благодаря встроенному в них искусственному интеллекту, становятся продолжением и дополнением нашего «Я», одновременно мягко и ненавязчиво ограничивая нашу мнимую свободу, наши целеполагания и действия.

Часто высказывается идея о виртуальной реальности, которая захватывает нашу обыденную реальность, замещает её собой, подменяя смыслы и значения вещей, людей и событий.

Виртуальная реальность выстраивается как отдельный мир, создаваемый с помощью технических средств. «Умные» вещи с одной стороны являются нашей собственной проекцией (с помощью голоса или нажатия кнопки мы на расстоянии приводим в действие некоторые механизмы), с другой стороны образуют особый мир IoT, существующий вполне автономно и нуждающийся в техническом сопровождении и обслуживании человеком. И в этом смысле поднимается ещё одна важная тема приоритета нужд человека и техники: техника создаётся человеком исключительно для его нужд, но чем более сложной становится техника как таковая, тем в большую зависимость от её надёжного функционирования попадает сам человек. Относительно работы IoT и самой возможности существования виртуальной реальности, наличие сети интернет, либо доступа к определённым сайтам и приложениям, оказывается критическим. Техническая инфраструктура оказывает решающее воздействие на всю социальную жизнь общества. А осуществление интернет соединения невозможно без электричества и т.д.

Зависимость человека от мира техники и взаимопроникновение мира реального и виртуального очень слабо осознаётся людьми. Технические артефакты необыкновенно органично и быстро входят в жизнь людей, дополняя и облегчая её, меняя наши привычки и стереотипы поведения. Объединение виртуального и реального миров иногда очень помогает нашему социальному «Я», а иногда приводит к раздвоению нашей личности. Интернет вещей расширяет границы нашего бытия, делая его подручным, а он, в свою очередь, начинает формировать наши запросы, желания и возможности.

Литература

Курбатов, В. И., Папа, О. М. (2023). Интернет вещей: основные концепции и тренды. *Гуманитарные, социально-экономические и общественные науки*. № 1. С. 48–54.

Журавлев А. Л., Нестик Т. А. (2019). Социально-психологические последствия внедрения новых технологий: перспективные направления исследований. *Психологический журнал*. Том 40. № 5. С. 35–47.

Ashton K. (2009). That «Internet of Things» Thing. *RFID Journal*. [Электронный ресурс]. URL: <http://www.rfidjournal.com/articles/pdf74986> (дата обращения: 13.05.2025).

Содержание

ПЛЕНАРНОЕ ЗАСЕДАНИЕ	3
<i>Барышников П. Н.</i>	3
Аналоговые вычисления и компьютерная метафора сознания: возвращение к истокам	
<i>Свириденко Д. И.</i>	8
Агентный искусственный интеллект и виртуальная реальность	
<i>Сторожук А. Ю, Зайкова А. С.</i>	12
Быть или казаться?	
Машина мысли:	15
эволюция компьютерной метафоры сознания в эпоху больших вычислений	
<i>Габов А. Н.</i>	14
Пифагорейское учение и архитектоника симуляции: неочевидные связи	

Дугушкин В. О., Кочергин В. И.

Применимость эпистемических
теорий в решении проблемы
галлюцинирования LLM

Ильин С. Е., Ильина А. В.

Причинно-следственные
отношения в исследованиях
естественного и искусственного
интеллекта

Кузнеченков А. А.

Модели сознания метасубъектов
третьей искусственной природы:
рекурсивный подход

Лихачева С. В., Донцов А. Д.

**Компьютерная метафора над
когнитивными стратегиями студентов:
последовательная и параллельная
архитектуры вычислительных систем**

Шипунова О. Д.

Машина мысли в суперинтеллектуальном обществе

Ярочкин Д. А.

Преодоление компьютерной
метафоры сознания через
аристотелевскую концепцию блага как
основы объяснимого и этически
выверенного ИИ

38

Проблемы феноменального сознания

38

Гаспаров И. Г.

Аргумент от представимости
«зомби» и категориальная ошибка

41

Горбачев М. Д.

Есть ли хорошие стратегии
определения феноменального
сознания?

<i>Гусев А. А.</i>	44
Панпсихизм, депсихологизация сознания и проблема комбинации	
<i>Десятко И. В.</i>	48
У натуралиста есть только один выбор: эмерджентизм или скептицизм	
<i>Морозов К. Е.</i>	52
Нормативная роль феноменального сознания	
<i>Павлов А. С.</i>	55
Мистерианство и аргумент многоуровневой абстракции	
<i>Сысоев М. С.</i>	58
Феноменальные состояния и феноменальные свойства	
ВЛИЯНИЕ ЦИФРОВИЗАЦИИ И ИНТЕЛЛЕКТУАЛЬНЫХ ТЕХНОЛОГИЙ НА ОБЩЕСТВО	62
<i>Баринов В. И.</i>	62
Процессы цифровизации и смартизации в дискурсе современной отечественной культуры	
<i>Вьшедко Г. В.</i>	65
Трансформация социальной коммуникации в ходе развития технологии искусственного интеллекта	

<i>Косенков А. Ю.</i>	69
Концепция четвертой промышленной революции: основные положения и теоретические принципы	
<i>Кузнецова А. В.</i>	73
Интеграция VR/AR в традиционное образование: развитие комплексных компетенций и адаптивных стратегий	
<i>Лихачева С. В., Попков А. А., Донцов А. В.</i>	77
Программное обеспечение для определения коэффициента стресса и фазы профессионального выгорания	
<i>Мельниченко Ю. А., Донцов А. В.</i>	81
Технологии и психика: как тревожность определяет технофобию и технофилию?	
<i>Некрасов Д. А.</i>	86
Структурные компоненты самосознания студентов, связанные с восприятием и формированием ограничений и самоограничений в цифровой среде	
<i>Скиба И. Р.</i>	89
Амбивалентные аспекты влияния интеллектуальных технологий на социум	
<i>Скрипкина Т. К.</i>	92
Позиционирование субъектности чат-бота с точки зрения поведенческих практик пользователей: попытка классификации	
<i>Ставровский И. К.</i>	96
Искусственный интеллект и проблема загрузки ценностей	

<i>Сущин М. А.</i>	99
Изменение образовательных практик в эпоху больших языковых моделей	
ГРАНИ ВИРТУАЛЬНОГО	103
	103
<i>Демина М. Г.</i>	
Сознание и цифровая реальность – грани диалектического противоречия	
<i>Лешкевич Т. Г.</i>	107
«Саморасширение» человека в контексте трансформаций антропосферы будущего	
<i>Лунев-Коробский О. А.</i>	110
О природной и нормативной реальности в неорационализме Р. Брассье	
<i>Миляева Е. Г.</i>	114
Цифровое посмертие: философские и технологические аспекты «загробной жизни» в цифровую эпоху	
<i>Новинская С. А.</i>	117
Современное искусство и новый антропогенез: от homo sapiens к homo digitalis?	
<i>Ускова Е. В.</i>	121
Виртуальность и реальность интернета вещей: технический, психологический и социально-философский аспекты	

УДК 303.01

ББК 87

А 437

Сборник издан по решению
Ученого совета Института философии и права СО РАН

Рецензент:

д-р филос. наук, проф. В. В. Целищев

Ответственные редакторы:

канд. филос. наук А. В. Хлебалин, канд. филос. наук А. С. Зайкова

А 437

СОЗНАНИЕ В ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ:
материалы Международной научной конференции / редкол.:
А. В. Хлебалин, А. С. Зайкова. – Новосибирск: Издательство
СО РАН, 2025. – 213 с.

ISBN 978-5-4437-0840-9

В сборнике опубликованы материалы Международной научной конференции «Сознание в виртуальной реальности», приуроченной к 75-летию теста Тьюринга. Книга рассчитана на специалистов в области философии сознания, когнитивных и социальных наук, а также на всех, кто стремится глубже понять природу сознания и то, как оно взаимодействует с виртуальными технологиями.

ISBN 978-5-4437-0840-9

© Институт философии и права
СО РАН, 2025

Научное издание

СОЗНАНИЕ В ВИРТУАЛЬНОЙ РЕАЛЬНОСТИ

Материалы Международной научной конференции

Редакционная коллегия

А. В. Хлебалин, А. С. Зайкова

Подписано в печать 00.10.2025 г.

Формат 60x84/16. Гарнитура Times New Roman

Уч.-изд. л. 00. Усл.-печ. л. 00

Тираж 00 экз. Заказ № 00

Издательство СО РАН