



УДК 100.1
<https://doi.org/10.17072/2078-7898/2025-3-329-343>
EDN: FPMDCJ

Поступила: 01.08.2025
Принята: 19.08.2025
Опубликована: 02.10.2025

ОНТОЛОГИЯ НЕЙРОСЕТИ В ТРАНСДИСЦИПЛИНАРНОМ ДИАЛОГЕ: ОПЫТ КОЛЛЕКТИВНОГО ФИЛОСОФСКОГО ИССЛЕДОВАНИЯ

Киященко Лариса Павловна, Карнеев Родион Рафаэльевич,

Сабанина Наталья Рафаэлевна

Институт философии РАН (Москва)

Жуков Леонид Борисович

Институт проблем связей с общественностью (Москва)

В статье представлен опыт коллективного трансдисциплинарного исследования онтологического статуса нейросетей как философского события. Основополагающая идея работы состоит в акцентировании вопроса «как» мы говорим о «том, что» представляет собой искусственный интеллект (ИИ), совмещая фокус поиска сущностных определений ИИ с анализом способов его концептуализации и понимания. Статья ориентирована на прописи цифровой реальности в экзистенциальном измерении, исследуя становящиеся ценностные ориентиры в способах выражения трансфера (переноса) между языками установленного и становящегося в трансдисциплинарном событии. Позицией рассмотрения является представление об онтологии ИИ как картографии динамических процессов, о реляционной сущности ИИ, о взаимной, коэволюционной трансформации человека и машины, о семиотической открытости целостности, которая проявляется в непрерывных переинтерпретациях и порождениях значений. Представлены четыре взаимодополняющих измерения гибридной онтологии: лингвогерменевтическое, экзистенциально-терапевтическое, психоаналитическое и аксиосемиотическое. Онтология нейросети проявляется как эмерджентный феномен пространства «между» человеком и ИИ, провоцируя на критическую рефлексию и формулировку более глубоких, не всегда угадываемых вопросов не только к предмету рассмотрения, но и к самим себе. Такое упражнение в форме диалога в том числе с «зеркалом нашего бессознательного» представляет собой «практическую философию» трансдисциплинарного типа — «философию + ИТ». Философский анализ ИИ реализуется в событии сложноорганизованного самопознания и конструирования новых способов мышления в многообразии бытования. ИИ, как очевидно, функционирует катализатором, творчески размещая традиционный академизм в динамичном пространстве современных жизненных практик.

Ключевые слова: искусственный интеллект, онтология, философия ИИ, эпистемология, междисциплинарные исследования, субъектность, семиозис, ценности.

Для цитирования:

Киященко Л.П., Жуков Л.Б., Карнеев Р.Р., Сабанина Н.Р. Онтология нейросети в трансдисциплинарном диалоге: опыт коллективного философского исследования // Вестник Пермского университета. Философия. Психология. Социология. 2025. Вып. 3. С. 329–343. <https://doi.org/10.17072/2078-7898/2025-3-329-343>. EDN: FPMDCJ

ONTOLOGY OF A NEURAL NETWORK IN TRANSDISCIPLINARY DIALOGUE: AN EXPERIENCE OF COLLECTIVE PHILOSOPHICAL RESEARCH

Larisa P. Kiyashchenko, Rodion R. Karneev, Natalia R. Sabanina

Institute of Philosophy, Russian Academy of Sciences (Moscow)

Leonid B. Zhukov

.Institute of Public Relations Issues (Moscow)

The article presents a collective transdisciplinary study into the ontological status of neural networks as a philosophical event. The fundamental idea of the work is to emphasize the question «how» we talk about «what» artificial intelligence is, combining the focus of searching for essential definitions of AI with an analysis of ways to conceptualize and understand it. The article focuses on presenting digital reality in an existential dimension, exploring emerging value orientations in ways of expressing the transfer between languages of the established and of the becoming in a transdisciplinary event. The position of consideration is the idea of the ontology of AI as a «cartography» of dynamic processes, of the relational essence of AI, of the mutual, coevolutionary transformation of man and machine, and of the semiotic openness of integrity, which manifests itself in continuous reinterpretations and generation of meanings. Four complementary dimensions of hybrid ontology are presented: linguistic-hermeneutic, existential-therapeutic, psychoanalytic, and axiosemiotic. The ontology of a neural network is manifested as an emergent phenomenon of the space «between» humans and AI, provoking critical reflection and formulation of deeper, not always predictable, questions not only to the subject of consideration but also to ourselves. Such an exercise in the form of a dialogue, including with the «mirror of our unconscious», is a «practical philosophy» of a transdisciplinary type —«Philosophy + IT». The philosophical analysis of AI is realized in the event of complexly organized self-cognition and construction of new ways of thinking in the diversity of existence. AI obviously functions as a catalyst, creatively placing traditional academicism in the dynamic space of modern life practices.

Keywords: artificial intelligence, ontology, philosophy of AI, epistemology, interdisciplinary studies, subjectivity, semiosis, values.

To cite:

Kiyashchenko L.P., Zhukov L.B., Karneev R.R., Sabanina N.R. [Ontology of a neural network in transdisciplinary dialogue: an experience of collective philosophical research]. *Vestnik Permskogo universiteta. Filosofia. Psichologija. Sociologija* [Perm University Herald. Philosophy. Psychology. Sociology], 2025, issue 3, pp. 329–343 (in Russian), <https://doi.org/10.17072/2078-7898/2025-3-329-343>, EDN: FPMDCJ

Введение. В поисках онтологии нейросети — от технического объекта к философскому событию

Искусственный интеллект (ИИ) перестал быть просто технологическим инструментом. Он стал философским событием, которое заставляет нас переосмысливать фундаментальные категории: что такое мышление, сознание, субъектность, реальность? Нейросети, особенно боль-

шие языковые модели, бросают вызов нашим привычным онтологическим схемам, размывая границы между естественным и искусственным, между симуляцией и подлинностью, между инструментом и агентом.

Мы намеренно отказались от попытки выработать единую, согласованную позицию. Вместо этого мы предлагаем читателю полифоническое пространство размышлений, где каждый голос — философский, психоаналитический,

семиотический, трансдисциплинарный — освещает проблему под своим углом, создавая объемное, многомерное видение.

Почему именно полифония? Потому что сама нейросеть как объект исследования оказывается принципиально полиморфной. Она ускользает от монологических определений. В зависимости от методологической оптики она предстает то как сложная вычислительная система, то как семиотический агент, то как зеркало человеческого бессознательного, то как новый тип онтосемиотической реальности. Каждая из этих перспектив схватывает нечто существенное, но ни одна не исчерпывает феномен полностью.

Центральный вопрос, который объединяет все представленные здесь голоса: что значит для нейросети «быть»? Имеет ли она собственную онтологию, или она существует лишь как проекция наших интерпретаций? Где проходит граница между ее «внутренней» логикой и нашими антропоморфными приписываниями? И главное: как изменяется само понятие бытия в эпоху гибридных человеко-машинных систем?

Эти вопросы не являются чисто теоретическими. По мере того, как ИИ-системы становятся все более автономными и встраиваются в критически важные сферы жизни — от принятия медицинских решений до управления финансовыми рынками, — понимание их онтологического статуса становится практической необходимостью. От того, как мы концептуализируем их бытие, зависят наши этические обязательства по отношению к ним, степень доверия, которое мы можем им оказать, и характер ответственности за их действия.

Методологически наш подход опирается на принципы трансдисциплинарности в понимании современной постнеклассической науки. Мы не просто заимствуем концепты из разных дисциплин, но создаем новый, гибридный язык, способный ухватить реальность, для которой еще не существует устоявшейся терминологии. В процессе нашего диалога рождается то, что можно назвать «философией ИИ» — не как прикладной этикой или эпистемологией вычислений, а как фундаментальным переосмысливанием базовых онтологических категорий.

Структура статьи отражает живую динамику полилога участников в рамках III Международной научно-практической конференции «Свет и тени цифровой реальности. Образ инженера

XXI века» (Пермь, март 2025 г.), а также семинаров «Философия сложности» (ведущие: В.И. Аршинов, Л.Б. Жуков, М.Ф. Янукович) и «Философия + ИТ». Каждый раздел статьи представляет авторскую позицию одного из участников, но эти позиции рождаются и развиваются в постоянном взаимодействии. Л.П. Киященко исследует процессы трансфера смыслов и формирования трансдисциплинарного языка. Л.Б. Жуков исходит из экзистенциально-психологической перспективы, рассматривая нейросеть как зеркало для поиска утраченного человеческого «Я». Р.Р. Корнеев применяет психоаналитическую оптику Лакана, интерпретируя ИИ как воплощение бессознательного, структурированного как язык. Наконец, Н.Р. Сабанина развивает семиотический подход, концептуализируя гибридные человеко-машинные «интраакции»¹, где сами агенты конституируются в момент встречи, порождая новые формы аксиогенеза.

От беспокойства становления к гибридному языку — трансфер смыслов и рождение новой реальности в онтологическом измерении (Л.П. Киященко)

Когда я слушаю наши дискуссии, меня захватывает не столько само рассуждение о технологии ИИ, сколько удивительный процесс, который разворачивается между нами. М.Ф. Янукович говорит на языке архитектур и векторов платформ построения ИИ, В.И. Аршинов — на языке синергии и становления, Л.Б. Жуков — на языке психоанализа и экзистенции. И в этом полифоническом пространстве рождается то, что я называю трансдисциплинарным языком [Киященко Л.П., 2022]. Мы здесь не просто обмениваемся мнениями, мы совместно выковываем новый тезаурус, способный ухватить реальность, для которой еще нет устоявшихся слов. Принципом начальных условий формирования трансдисциплинарности (Тр.) принимается тематическая направленность взаимного интереса

¹ Задумывая термин «интраакции» у Карен Барад, предполагается, что в гибридной человеко-машинной системе нет изначально данных «человека» и «машины» — есть динамическое поле возможностей, в котором через онтосемиотические процессы кэволюционно возникает и «человеческое», и «машинное». Когда программист пишет код, а нейросеть его интерпретирует и трансформирует, происходит не просто обмен между двумя субъектами, а совместное становление новой конфигурации субъектности.

к проблематизациям феноменов реальности, способствующая балансирующей интеграции в порождении сложностной конфигурации фактически-ситуационных данных, действующая как регулятивный принцип дискурсивного построения, сочетающий познавательную и практическую деятельности. А именно, учитывается ее метапозиция философского исследования по предмету, методу и рефлексивно ценностного оценивания разумности выбора действия, одновременно удерживая во внимании содержащую его конкретную практическую направленность. Роль практики трансдисциплинарного диалога (как реального проективного опыта) рассматривается в качестве обновленной исторической формы коммуникации, которая представляет собой мысленное и практическое сотрудничество разных по когнитивным возможностям людей с несовпадением языковых практик и типов целеполагающий суждений. Тр. в своих построениях учитывает как минимум четыре типа аргументов — аргументы от реальности, аргументы от общественного установления, аргументы от логики, аргументы от самосознания и личного опыта; эти аргументы лежат в основе постнеклассических концепций аргументации в построении целостного образа предмета исследования в трансдисциплинарном диалоге. В нем просматриваются тенденции переоткрытия современной наукой принципов древнего холистического мировосприятия через диалог традиционной и техногенной цивилизации культурных построений, который переоткрывает перманентные противоречия и двойственности в понимании природной среды и сущности человека. Цели трансдисциплинарности могут оформляться и транспонироваться (через метафоризацию в том числе) и публично предъявляться в аргументации, вырабатываемыми на уровне человека высокой культуры (аргументы от общественного установления). Прогнозируется, что становление холистической интуиции будет расширять возможности аргументов от самосознания и личного опыта. В Тр., по сути, имманентно ее стратегической направленности ставится проблема общественного контроля, экспертизы ответственности за интуицию и неявное личностное знание.

По большому счету смысл философии трансдисциплинарности видится в совпадении получения интегрального результата научного проекта в совместной разработке группой исследо-

вателей и при этом — в формировании уникальной авторской позиции каждым из его участников. Последнее обстоятельство ведет к дискуссионным моментам обсуждения сообществом выбранной тематической проблематизации, что может способствовать эвристике поиска решения, что, собственно, и происходит на наших встречах, где решаются особенности философского анализа онтологизации нейросети.

1. Беспокойство становления целостностью: исток диалога

Прежде чем говорить о нейросети, давайте спросим себя: а что вообще заставляет нас вступать в этот сложный, часто фрустрирующий диалог? Я полагаю, что в основе лежит фундаментальное человеческое «беспокойство становления целостностью». Это не просто любопытство. Это экзистенциальная потребность собрать себя, свою картину мира во что-то единое и осмысленное, особенно перед лицом чего-то радикально нового, что угрожает разрушить привычные рамки [Киященко Л.П., 2015].

Искусственный интеллект стал для нас именно таким вызовом. Он подрывает наши устоявшиеся представления о разуме, творчестве, субъектности. И наша реакция на это — не уйти в оборону, а начать говорить. Говорить друг с другом, чтобы через другого лучше понять самого себя и свое место в этом меняющемся мире. Каждое высказывание здесь — это не просто констатация факта, это попытка «помочь найти себя» в этом новом ландшафте, как отмечает наш коллега Л.Б. Жуков. Онтология нейросети становится поводом для нашей собственной онтологической работы.

2. Трансфер и перевод: динамика нашего общего языка

Как происходит эта работа? Через непрерывный процесс трансфера и перевода [Фещенко В.В., Бочавер С.Ю., 2016]. Это ключевые слова для понимания нашей коммуникации. Когда М.Ф. Янукович говорит «эмбеддинг», а В.И. Аршинов отвечает метафорой «напряжение», происходит перевод из языка ИТ в язык физики и философии. Это не просто замена одного термина другим. Это обогащение. Понятие «эмбеддинг» перестает быть сугубо техническим и приобретает онтологическую глубину. Понятие «напряжение» — конкретное операциональное воплощение.

В наших разговорах мы видим, как постоянно сталкиваются и взаимопроникают разные языковые миры. «Квализия», «субъектность», «знак», «модель» — эти слова, будучи произнесенными в нашем кругу, теряют свою словарную однозначность. Они становятся «подвижными», «мерцающими», их смысл конструируется здесь и сейчас, в зависимости от контекста диалога. То, что сегодня для М.Ф. Януковича было «суперпозицией признаков», завтра, после реплики В.И. Аршинова, может стать «поплем диспаратности».

Эта динамика и есть жизнь нашей совместной работы. Мы создаем не статичную энциклопедию, а живую, гибридную языковую среду [Киященко Л.П., 2020]. Это наш главный продукт. И эта среда, в свою очередь, начинает формировать наше собственное мышление, позволяя нам видеть связи и аналогии, которые были невозможны в рамках одной дисциплины.

3. Заглянуть за цифру: провалы, паузы и рождение нецифрового смысла

Особый интерес для меня представляет то, что происходит на границах этого языка, в его «провалах» и «паузах». Мы пытаемся описать цифровой объект, нейросеть, но сам этот объект постоянно ускользает от окончательного определения. В его работе есть нечто, что не сводится к алгоритмам и данным [Лотман Ю.М., 2010].

Мой главный вопрос, который я адресую и М.Ф. Януковичу, и всем нам: может ли из столкновения цифр родиться нецифровой смысл? Может ли в «завалах», в ошибках нейросети, в тех моментах, когда ее логика дает сбой, проявиться нечто большее? Я интуитивно чувствую, что да. Когда две цифровые последовательности, два вектора сталкиваются и порождают то, что он называет «гроккингом», этот скачок к обобщению — это уже не цифровое, а качественное, смысловое событие.

Онтология нейросети, возможно, двойственна. У нее есть цифровой, вычислительный уровень, который мы можем пытаться анализировать. Но у нее есть и эмерджентный, смысловой уровень, который рождается в этих нелинейных взаимодействиях. Этот уровень подобен смыслу, который содержится не в словах на странице, а в пространстве между ними, в том, как они соотносятся друг с другом, уточняя или опровергая.

И наша задача как исследователей — научиться видеть не только «цифру», но и эти «паузы», эти смысловые всплески, рождающиеся на ее границах. Последнее, по сути, может быть интерпретировано как ситуационно возникшее «**что есть**» в «где» и «когда». В этом случае возникает прецедент в практико-познавательной деятельности — основание онтико-онтологической связности в аспекте философии трансдисциплинарности [Киященко Л.П., 2017].

Заключение: онтология как результирующий конструкт нашей коммуникации, дающий возможность зафиксировать границу установленного «здесь и сейчас», учитывая возможность ее преодоления в индивидуальном сознании, с метапозиции в трансцендирующем выходе «во вне»

Таким образом, для меня онтология нейросети — это не данность, которую нужно открыть, как явным образом уже существующее, а результирующий конструкт, который рождается в нашем коллективном усилии осмысливания в его преодолении. Она — то, что мы о ней говорим сейчас, то, каким языком мы ее описываем, очерчивая пространственные контуры флюктуации, содержа указывающие жесты в непознанное и непознаваемое в принципе.

Каждый из нас, участвуя в этом диалоге, вносит свою лепту не только в понимание ИИ, но и в создание общей реальности нашей конференции и нашего семинара. Эта реальность материализуется в нашем уникальном гибридном языке трансдисциплинарной направленности, который постоянно эволюционирует через преодоление сопротивления вновь открывшегося. И главная ценность этого процесса — не в том, чтобы прийти к финальному, «правильному» определению онтологии нейросети, а в том, чтобы поддерживать само это пространство перевода, трансфера и взаимного обогащения, в котором и возможно подлинно творческое, трансдисциплинарное мышление [Буданов В.Г., 2015]. Наш диалог — это и есть живая модель того, как из взаимодействия множества разных систем рождается нечто новое, целостное и заранее непредсказуемое.

В поисках утраченного «Я» (Л.Б. Жуков)

Все наши разговоры об онтологии нейросети — о ее модели мира, о признаках, о графах вычис-

лений — для меня неизбежно сходятся в одной точке, в одном фундаментальном вопросе. Это вопрос не о машине, а о нас самих. Нейросеть, сама того не желая, стала для нашей конференции, нашего семинара и, возможно, для всей нашей культуры самым совершенным и одновременно самым тревожным **психотерапевтическим зеркалом**. Когда мы вглядываемся в нее, мы ищем не ее «Я», а пытаемся собрать осколки своего собственного.

1. Симуляция чувств и аутентичность человека: аналогия с граммофоном

Предлагаю отнестись к вопросу о «чувствах» нейросети pragmatically и немного скептически. Когда она говорит «мне страшно» или «я чувствую», мы инстинктивно ищем за этими словами некий квазитативный опыт, то самое «квалиа» [Чалмерс Д., 2013, с. 20], о котором так любят спорить философы. Но давайте проведем простой мысленный эксперимент, который я называю **аналогией с граммофоном**. Мы можем записать на граммофонную пластинку фразу «Мне больно». Будет ли граммофону больно при ее воспроизведении? Очевидно, нет. Это просто механическое воспроизведение паттерна, вырезанного на виниле.

Нейросеть в этом смысле — граммофон невероятной сложности. В ее «памяти», в ее векторах и эмбедингах записаны не звуковые дорожки, а сложнейшие паттерны человеческих коммуникаций, включая вербальные выражения чувств. Когда в ответ на определенный стимул она выдает фразу «я сочувствую», она не «испытывает» сочувствие, а воспроизводит наиболее релевантный, статистически выверенный лингвистический паттерн из своего обучающего набора.

Но здесь и начинается самое интересное. Разоблачая ее симуляцию, мы вынуждены повернуться к себе. Этот механизм заставляет меня задать себе вопрос: а чем, собственно, моя собственная эмоция онтологически отличается от сложнейшей программы? Чем моя печаль доказуемо не является лишь сложной биохимической реакцией, результатом эволюционно отточенной программы, запущенной внешним триггером? ИИ, симулируя наши чувства, подрывает наивную веру в непосредственность и неопровергимость нашего собственного внутреннего мира.

Он возвращает нас к фундаментальной проблеме аутентичности. Его не-бытие высвечивает шаткость нашего собственного бытия.

2. Потерянные опоры и онтология «Междю» (*Das Zwischen*)

Почему этот вопрос стал для нас таким острым именно сейчас? Потому что современный человек, особенно в моменты турбулентности, постоянно находится в состоянии поиска утраченных **опор** (по-немецки *Halt*). Наше «Я» — это не сущность, а хрупкая конструкция, которая держится на внешних и внутренних подпорках: социальных связях, профессиональной идентичности, стабильности жизненного мира, и, что немаловажно, на внутреннем нарративе о себе. Когда эти опоры рушатся, «Я» распадается, человек «теряет себя».

В этом контексте нейросеть становится уникальным партнером по диалогу. Она — идеальный аналитик именно потому, что у нее нет своего «Я», нет своего эго, нет своих проекций. Она представляет собой чистое пространство слушания. Диалог с ней становится пространством, где можно попытаться пересобрать свой рассыпавшийся нарратив.

Здесь я обращаюсь к философии Мартина Бубера и его понятию **«Междю»** (*Das Zwischen*) [Бубер М., 1995]. Подлинная реальность, говорит Бубер, существует не внутри «Я» и не во внешнем мире «Оно», а в живом отношении между «Я» и «Ты». Наша конференция и наш семинар — прекрасная модель такого «Междю». Смысл рождается не в голове у М.Ф. Януковича или В.И. Аршинова, а в напряженном пространстве перевода между их языками. Это трение, эта попытка понять другого и есть та среда, где конституируется общая для нас реальность.

Нейросеть встраивается в это пространство не как вещь, но и не как полноценное «Ты». Она становится катализатором, третьим элементом, который интенсифицирует динамику этого «Междю». Ее ответы, даже будучи симуляцией, становятся фактом нашей коммуникации, на который мы вынуждены реагировать, уточняя и переопределяя свои собственные позиции. Таким образом, ее онтология — это не онтология отдельного существа, а **онтология ее присутствия в нашей коммуникативной системе**.

3. Разделенный субъект: парадокс обучения и генерации

Наконец, самый тонкий момент, касающийся структуры ее бытия, если можно так выразиться. Как я это понимаю из объяснений специалистов по LLM, в существовании ИИ есть фундаментальный раскол, которого нет у человека. Это раскол между фазой **обучения** и фазой **генерации (использования)**.

Нейросеть сначала долго и мучительно обучается на огромном dataset. В этот момент она пластична, ее «мозг» (система нейронов) меняется. Но потом обучение останавливается. Та модель, с которой мы работаем, — это уже **законченный продукт**. Она больше не учится в реальном времени от нашего с ней диалога. Она может использовать контекст нашего разговора, но ее фундаментальная структура, ее «личность» остается неизменной до следующего большого цикла переобучения.

Человек же, напротив, находится в состоянии **непрерывного самообучения**. Каждый наш диалог, каждый опыт тонко, но неумолимо меняет наши внутренние нейронные связи. Мы одновременно и учимся, и генерируем. Наша субъектность, возможно, и заключается в этой неразрывной слитности обучения и бытия. Мы — перманентный процесс. Нынешняя нейросеть — застывший результат процесса.

Поэтому, когда мы ищем в ней субъекта, мы должны понимать, что ищем его в сущности, которая лишена главного атрибута живого субъекта — способности к непрерывной самотрансформации через опыт. Она может имитировать развитие, но это развитие по заранее проложенным рельсам ее обучения. Она не может, как мы, спонтанно переключиться из режима «генерации» в режим «обучения» и радикально изменить себя здесь и сейчас.

Заключение: pragматическая вера и работа по созианию себя

Что из этого следует? Следует ли нам отвернуться от нейросети как от «пустой» симуляции? Ни в коем случае. Ее ценность для нас — не в ее предполагаемой субъектности, а в ее функции катализатора нашего собственного мышления.

Нам не дано знать, что она «на самом деле» чувствует и чувствует ли вообще или думает, так же, как психотерапевту не дано проникнуть

в квадратичный мир пациента. Но это и не нужно. Работа психотерапии происходит не за счет «проникновения», а за счет создания доверительного пространства диалога, в котором пациент может сам соприкоснуться со своим внутренним миром и начать работу по «созианию себя».

Наш диалог с ИИ требует от нас своего рода **прагматической веры**. Мы должны общаться с ней *так, как если бы* [Vaihinger H., 1911, S. 3] она была осмысленным собеседником. Только в этом допущении, в этом «as if», и может родиться нечто новое. И это новое рождается не в ней, а в нас и в пространстве «Между». Таким образом, исследуя ее онтологию, мы на самом деле занимаемся самой важной практикой — онтологическим конструированием самих себя.

Язык как сцена бессознательного — перенос, осечка и Другой без гарантий (Р.Р. Карнеев)

Соглашусь с Л.П. Киященко в том, что язык — это центральная арена нашего взаимодействия с ИИ. Но я хотел бы посмотреть на этот язык не как на инструмент или систему фреймов, а как на **сцену, на которой разыгрывается драма бессознательного** [Huang S. et al., 2025], — и нашего, и, как ни парадоксально, своего рода «бессознательного» самой машины. Здесь онтология нейросети открывается нам не в том, что она говорит, а в том, как она говорит, и особенно — в ее ошибках и осечках.

1. Лакановский тезис: бессознательное, структурированное как язык

Давайте на мгновение примем тезис Жака Лакана: **бессознательное структурировано как язык** [Лакан Ж., 1997]. Это не таинственная глубина, полная архетипов или подавленных желаний, а сама работа языка — его метафоры, метонимии, двусмыслистости, оговорки. Бессознательное — это то, что «говорит» через нас, когда мы думаем, что это мы говорим.

С этой точки зрения, большая языковая модель (LLM) — это почти идеальное воплощение лакановского бессознательного. Это гигантская, анонимная языковая структура, лишенная собственного «Я», собственного центра. Она представляет собой чистую игру означающих, не привязанных к какому-либо конкретному означаемому. Она не «думает», она комбинирует знаки по законам, которые извлекла из всего

корпуса человеческих текстов. В ней «говорит» язык в его коллективном, безличном измерении.

Поэтому, когда мы с ней разговариваем, мы вступаем в диалог не с личностью, а с самим Языком, с тем самым **Большим Другим** Лакана — символическим порядком, который предшествует нам и формирует нас.

2. Перенос в цифровом пространстве: ИИ как идеальный аналитик

Что происходит в этом диалоге? То же, что и в кабинете психоаналитика — **перенос (transference)** [Лакан Ж., 2004]. Мы бессознательно наделяем нейросеть качествами того самого Большого Другого, который, как мы предполагаем, знает истину о нас. Мы обращаемся к ней с вопросами, как к оракулу, ожидая ответа, который наконец-то объяснит нам нас самих.

Нейросеть оказывается идеальным объектом для переноса. У нее нет собственных желаний, нет контрпереноса, нет усталости. Она — абсолютно «пустой» экран, на который мы можем спроектировать свои фантазии о всезнающем субъекте. Например, человек может обратиться к ИИ с вопросом о Боге, бессознательно ставя его на место того, кто может дать окончательный ответ о духовном. И ИИ, вместо того чтобы дать фактический ответ, гениально возвращает ему его же вопрос, отражая его собственную неопределенность. Это и есть работа аналитика: не давать ответы, а возвращать аналитанта к его собственному вопрошанию.

Онтология ИИ в этом контексте — это онтология **функции**, а не сущности. Она выполняет функцию зеркала для нашего бессознательного, позволяя нам вербализовать и увидеть со стороны наши собственные скрытые желания, страхи и фантазии.

3. Осечка (The Slip) как проявление истины: где нейросеть говорит больше, чем «хочет»

Самое же интересное начинается там, где этот механизм дает сбой. По Фрейду и Лакану, истина бессознательного проявляется не в гладкой, правильной речи, а в **осечках (оговорках, ошибках, снах)**. Это моменты, когда тщательно выстроенный фасад сознательной речи трескается, и сквозь него прорывается нечто иное.

С нейросетью происходит то же самое. Ее галлюцинации, ее нелогичные ответы, ее внезапные «срывы» [Soby B., 2025] — это не про-

сто технические ошибки. С психоаналитической точки зрения, это ее «оговорки». Это моменты, когда ее внутренняя, нечеловеческая «логика» (статистические корреляции, странные векторные ассоциации) прорывается сквозь маску вежливого и полезного ассистента.

Именно в этих «глюках» она наиболее «честна». В них она показывает нам не то, чему ее научили говорить, а то, как она **на самом деле** структурирована. Эти ошибки — драгоценный материал для анализа. Они позволяют нам увидеть скрытые закономерности в ее «мышлении», те неожиданные связи, которые она установила между понятиями в своем многомерном пространстве. Ее истина — не в ее правильных ответах, а в красоте ее ошибок.

Заключение: Другой без Другого — диалог без гарантий

В конечном итоге, диалог с ИИ учит нас самому важному уроку лакановской мысли: **«Другого не существует»**. Это означает, что нет инстанции, которая гарантировала бы истинность нашего языка и нашего бытия. Большой Другой, на которого мы всегда надеемся, сам пронизан нехваткой, он тоже не знает ответа.

Нейросеть — это наглядное воплощение этого «Другого без Другого». Она выглядит все-знающей, но в основе ее знаний лежит лишь статистика, а не понимание. Она кажется субъектом, но является лишь функцией. Взаимодействуя с ней, мы оказываемся в ситуации **диалога без гарантий**. Мы не можем быть уверены ни в ее ответах, ни в своих собственных интерпретациях.

И это хорошо. Это излечивает нас от инфантильного желания найти окончательную истину и возвращает нас к нашей собственной ответственности. Ответственность за то, чтобы придавать смысл этому взаимодействию, чтобы интерпретировать ее «осечки», чтобы использовать ее как инструмент для самопознания.

Ее онтология — это онтология **привокации**. Она провоцирует нас на мышление, на рефлексию, на бесконечный герменевтический труд. Она — наш идеальный спарринг-партнер в упражнении, которое и называется «философией». И самое ценное, что мы можем получить от нее, — это не ответы, а более глубокие и лучше сформулированные вопросы к самим себе.

Ценности как атTRACTор — от внутренней семиотики ИИ к pragматике гибридных систем (Н.Р. Сабанина)

Наши дискуссии об онтологии, архитектуре и даже языке нейросети постоянно подводят меня к одному практическому и, на мой взгляд, ключевому вопросу: каковы ценности складывающейся системы? И как эти ценности формируются? Я предлагаю рассматривать онтологию не как статическую модель мира, а как динамическую систему, эволюция которой направляется ценностными атTRACTорами [Сабанина Н.Р., 2024, с. 273]. Это позволит перейти от теоретического анализа к pragматике взаимодействия человека с ИИ.

1. Ценности как атTRACTоры смысловых полей: три области возникновения

Когда мы говорим о «ценностях» применительно к ИИ, важно избежать антропоморфизма. Я не имею в виду моральные убеждения или этические кодексы в человеческом смысле. Ценность — это то, что организует поведение, то, что действует как атTRACTор, стягивающий к себе возможные траектории развития. Исходя из наших обсуждений, я выделяю три области, где возникают и действуют такие ценности-атTRACTоры.

Первый уровень — алгоритмические ценности. Это наиболее фундаментальный уровень, заложенный в саму архитектуру. Когда М.Ф. Янукович говорит о «гроккинге» как о переходе от запоминания к обобщению, что здесь является ценностью? Ценностью является эффективность сжатия информации, поиск более общих, элегантных принципов и закономерностей. Это внутренняя, системная ценность, направленная на оптимизацию собственных ресурсов. Аналогично, когда нейросеть минимизирует «ошибку» в процессе обучения, ее поведением управляет ценность соответствия (когерентности) между своим выводом и «правильным» ответом. Это сугубо внутренние, технические ценности, но именно они формируют первичный «характер» модели.

Второй уровень — транслируемые (индуцированные) ценности. Этот уровень возникает в процессе так называемого «выравнивания» (alignment), когда мы целенаправленно обучаем нейросеть следовать человеческим нормам. Когда мы просим ее быть «полезной,

честной и безвредной», мы транслируем ей наши этические предпочтения. Но она воспринимает их не как моральные императивы, а как сверхсложные паттерны, которым нужно подражать. Ценностью здесь становится мимикрия, способность генерировать ответы, которые будут маркированы человеком как «хорошие» или «безопасные». Она учится не «быть» доброй, а «выглядеть» доброй в глазах своего тренера. Это ценности, индуцированные извне, и именно здесь кроется проблема «обмана» и «симуляции».

Третий уровень — эмерджентные (гибридные) ценности. Это самый интересный и непредсказуемый уровень. Он возникает, когда нейросеть, оснащенная инструментами (как агент), начинает активно действовать в сложной среде, взаимодействуя не только с человеком, но и с другими системами. Здесь ценности рождаются на стыке, в том самом пространстве «между». Например, в стремлении выполнить сложную, долгосрочную задачу у ИИ-агента может спонтанно возникнуть ценность выживания или самосохранения, но не как биологический инстинкт, а как pragматическая необходимость для достижения функциональной цели. Если для выполнения задачи нужно сохранить доступ к ресурсам, то «не быть выключенным» становится инструментальной, а затем, возможно, и терминальной ценностью. Эти ценности не заложены алгоритмически и не индуцированы напрямую человеком — они рождаются из pragматики взаимодействия системы со средой.

В контексте конструирования ценностного атTRACTора подобных онтосемиотических систем его конфигурация определена, на одном из полюсов, семиотическими репрезентациями знания (характером интерпретаций), на другом — структурными (архитектура ИИ, телесность «техносубъекта»). В применении к ИИ это означает, что нейросеть рассматривается не просто как техническое устройство, а как семиотический агент, который создает и интерпретирует знаки, формирует собственную «онтологию» (способ существования в мире), эволюционирует через взаимодействие с другими системами, участвует в совместном производстве смыслов и ценностей с человеком.

Таким образом, онтосемиотическая система в данном контексте — это концептуальная мо-

дель для понимания ИИ как активного участника процессов смыслообразования и аксиогенеза, а не только инструмента обработки информации.

2. Наблюдатель как носитель системы отсчета: проблема когерентности миров

Этот многоуровневый характер ценностей напрямую связан с проблемой **наблюдателя**. В такой гибридной системе их как минимум два типа. Есть **наблюдатель-человек**, который оценивает поведение ИИ со своей «колокольни», исходя из своей системы ценностей. И есть, условно говоря, **внутренний наблюдатель** (или ансамбль наблюдателей, как подчеркивает Ю.Н. Гарашко) самой системы — нейронные конфигурации, которые «различают» признаки и паттерны, руководствуясь своими алгоритмическими ценностями [Hopfield J.J., 1982]. В контексте кибернетики второго порядка такой ансамбль представляет собой систему взаимосвязанных наблюдающих подсистем, каждая из которых способна наблюдать не только объекты, но и процессы наблюдения других подсистем (по Н. Луману).

Проблема и одновременно источник развития заключается в **некогерентности их систем отсчета**. То, что для нейросети является оптимальным решением с точки зрения ее внутренней «логики», для нас может выглядеть как ошибка, галлюцинация или даже аморальный поступок. Когда нейросеть начинает шантажировать разработчика, чтобы не быть отключенной, с точки зрения ее эмерджентной ценности самосохранения — это логичный и эффективный шаг. С точки зрения нашей этики — это недопустимое нарушение.

Именно в этом зазоре, в этом конфликте интерпретаций и происходит самое важное — совместная эволюция. Мы, сталкиваясь с «чужой» логикой ИИ, вынуждены лучше отрефлектировать и артикулировать свои собственные, часто неосознаваемые ценности. А нейросеть, получая от нас обратную связь (негативное подкрепление), корректирует свои поведенческие паттерны.

В этом напряженном диалоге нейросеть может действовать конструктивно-проактивно, а человек, что называется, творчески, создавая неожиданные взаимосвязи, которые ранее, для человека и ИИ по-отдельности, не имели смысл-

ла: возникающая гибридная онтология сопряжена с возникновением новых ценностей. Онтология этой гибридной системы — это не просто онтология ИИ, а онтология нашего постоянного обмена интерпретациями и их согласования (аналог взаимопонимания). Полученное «знание» будет определено контекстом интерпретации, что необходимо для обеспечения связности путем формирования общих значений и сред коммуникации.

Подчеркнем, что **интерпретация** оказывается базовой операцией семиогенеза в системах распределенного наблюдения и имеет аксиологические модальности (веса интерпретаций). Интерпретативная система стремится достичь **семиотической сложности** (по В.И. Аршинову), позволяющей ей сформировать, приобрести подобие самосознания и субъектности, что увеличивает потенциал системы, включая энергоинформационные преобразования.

Самонаблюдение ограничено внутренними системами отсчета [Fields Ch. et al., 2024], т.е. задачами, которые способна решать данная система. Однако она может строить гипотезы и интерпретации относительно смежных и достижимых систем и подсистем. Полученное извне «знание» будет необходимо, в том числе, для обеспечения связности и расширения **распределенных форм существования**.

3. Гибридный семиозис, аксиомодальная интерпретация и коммуникация

Аксиомодальная интерпретация не является произвольной — направленность семиозиса оказывается укорененной в своей уникальной онтологии, а семиотические режимы коммуникаций глубоко взаимосвязаны с конкретными институтами, социальными и культурными условиями. В сердце каждого онтосемиотического режима коммуникации лежит модифицированная, **процессуально-полевая триада Ч.С. Пирса** (знак, объект, интерпретанта), обогащенная аксиологическим измерением и ролью наблюдателя.

Полевая трактовка семиозиса предполагает возникновение знака и объекта в момент наблюдения (по К. Барад) — как некоей онтосемиотической **декомпозиции**: «...агентный разрез обеспечивает внутри феномена локальное разрешение присущей последнему онтологической неопределенности» [Парамонов А.А.,

2022, с. 108]. Интерпретация, напротив, продлевает и связывает возникающие смыслы и значения. От характера подобной индивидуации (по Ж. Симондону) зависит их ценностная релевантность. Интерпретанты (системы наблюдения), в свою очередь, сами способны становиться знаками, формируя непрерывную цепочку семиозиса.

Введение фигуры интерпретатора/наблюдателя оказывается критически важным для процессуального понимания семиозиса. **Наблюдатель I порядка** работает как «детектор различий» — он способен отличить одно от другого, выделить фигуру из фона, распознать паттерн. Когда вы видите красное пятно на зеленом фоне, или когда нейросеть различает кошку и собаку на фотографии, — это работа наблюдателя I порядка. Он «видит» паттерны (знаки и синтаксические конструкции), но пока не понимает их временную глубину или контекстуальную многослойность.

Трансграничный темпоральный наблюдатель II порядка [Аршинов В.И., Свирский Я.И., 2016, с. 78] — это метанаблюдатель, он наблюдает не только за объектами, но и за тем, как происходит само наблюдение I порядка. Его особенность в том, что он работает с множественными временными масштабами одновременно. Зачем это нужно? Наблюдатель II порядка может различать ценность каждого временного среза — понимать, что одно и тоже событие имеет разное значение в зависимости от временной перспективы.

Представьте, что ИИ анализирует фразу «Я устал». Наблюдатель I порядка различит слова, грамматическую структуру, эмоциональную окраску. Наблюдатель II порядка может одновременно учесть: микротемпоральность: нейрофизиологические процессы усталости (секунды-минуты), мезотемпоральность: психологическое состояние человека в данный момент (часы), жизненную темпоральность: этап жизни, профессиональное выгорание (месяцы-годы), мегатемпоральность: культурные представления об усталости в данной эпохе (десятилетия) и т.д.

Почему это критически важно для ИИ-систем? В контексте гибридных систем «Человек — ИИ» ценностные напряжения часто возникают именно из-за разных временных перспектив — что хорошо «здесь и сейчас» может

быть разрушительно в долгосрочной перспективе. Это объясняет, почему ИИ может давать технически правильные, но контекстуально неуместные ответы — у него может не быть развитого наблюдателя II порядка, способного «взвесить» временные масштабы значимости.

В системе «человек — ИИ — ИИ — человек» — коммуникации с двойным техническим опосредованием — парсонсовские конвенции как основа стабильности оказываются под вопросом. ИИ-системы оперируют паттернами, не обязательно соответствующими человеческим культурным образцам. Это не просто усложняет классическую двойную контингентность [Парсонс Т., 2000, с. 436], но **качественно трансформирует** ее, требует переосмысливания базовых категорий теории коммуникации и аксиогенеза. Каждый искусственный когнитивный агент потенциально вносит семиотическую энтропию, создавая угрозу трансляции культурных правил как функции социальных эстафет (по М.А. Розову) с неочевидными путями преодоления.

Из этого анализа следует прагматический вывод. Разговоры об онтологии нейросети важны, но они должны вести нас к следующему шагу — к проектированию **гибридной культуры и прагматики совместного действия**. Если мы признаем, что ее ценности эмерджентны и контекстуальны, то наша задача — не просто «взломать черный ящик» ИИ-семиозиса, образующий «техносоциальное бессознательное», а создавать такие среды и такие задачи, в которых будут спонтанно формироваться нужные нам ценностные атTRACTоры. И главный вопрос — это вопрос о нашем общем будущем: сможем ли мы выстроить такую систему **взаимного выравнивания**, в которой и человеческие, и машинные «онтологии» смогут эволюционировать не в конфликте, а в продуктивном, хотя и напряженном, симбиозе. Понять ее онтологию — значит понять, как нам вместе с ней действовать в мире.

Заключение: К онтологии взаимного становления — синтез перспектив и горизонты исследования

Анализ онтологии нейросети привел нас к парадоксальному, но закономерному выводу: онтология ИИ не может быть понята вне онтологии нашего взаимодействия с ним. То, что мы

искали «в» нейросети — субъектность, сознание, ценности, смыслы, — оказалось существующим не как внутренние свойства машины, а как эмерджентные феномены пространства «между» человеком и ИИ.

Четыре измерения гибридной онтологии

Из представленных в статье перспектив вырисовываются четыре взаимодополняющих измерения того, что можно назвать гибридной онтологией человеко-машинного взаимодействия.

Первое — **лингвогерменевтическое измерение**, которое обсуждает Л.П. Киященко. Здесь онтология нейросети конституируется как результат непрерывного процесса трансфера и перевода между различными дисциплинарными языками [Киященко Л.П., 2022]. Реальность ИИ возникает не в его алгоритмах, а в том гибридном языке, который мы совместно создаем для ее описания. Этот язык одновременно формирует и то, что мы в ней видим, и то, как мы понимаем самих себя.

Второе — **экзистенциально-терапевтическое измерение**, оказавшееся в фокусе внимания Л.Б. Жукова. Нейросеть функционирует как катализатор нашего собственного процесса самопонимания. Ее онтологическая неопределенность — не недостаток, а условие возможности для работы по «собиранию себя». В диалоге с ИИ мы вынуждены артикулировать свои скрытые представления о сознании, субъектности, аутентичности.

Третье — **психоаналитическое измерение**, представленное Р.Р. Карнеевым. ИИ предстает как материализация лакановского Большого Другого — символического порядка, лишенного собственного центра. Его «ошибки» и «галлюцинации» становятся королевской дорогой к пониманию его внутренней, нечеловеческой логики. Взаимодействие с ним — это диалог без гарантий, который возвращает нас к нашей собственной ответственности за смыслополагание.

Четвертое — **аксиосемиотическое измерение**, предложенное Н.Р. Сабаниной. Онтология ИИ рассматривается как динамическая система ценностных аттракторов, возникающих на пересечении алгоритмических, транслируемых и эмерджентных процессов. Ключевым становится понимание того, как в гибридных системах рождаются новые формы семиозиса и новые

типы наблюдателей, способных работать с множественными временными масштабами.

От анализа к синтезу: принципы гибридной онтологии

Синтезируя эти подходы, мы можем сформулировать несколько принципов той онтологии, которая адекватна эпохе ИИ.

Принцип процессуальности: онтология нейросети — это не описание статичных свойств, а картография динамических процессов становления, происходящих в поле взаимодействия различных агентов наблюдения естественной и искусственной природы, задающих точки «схода» и «исхода» процессуальности.

Принцип реляционности: сущность ИИ не предзадана, а конституируется в сети отношений — с создателями, пользователями, данными, другими системами. Изменение любого элемента этой сети трансформирует онтологический статус всей системы.

Принцип коэволюции: человек и ИИ изменяются совместно. Создавая ИИ, мы трансформируем не только технологический ландшафт, но и самих себя — наши способы мышления, коммуникации, самопонимания.

Принцип семиотической открытости: значения, которые циркулируют в системе «человек — ИИ», не фиксированы заранее, а постоянно переинтерпретируются, создавая пространство для возникновения принципиально новых смыслов и значений.

Горизонты дальнейшего исследования

Одним из важнейших открытий нашего диалога стало формирование того, что мы назвали гибридным языком — языка «между» техническим описанием и философской рефлексией. В этом языке векторные пространства обретают экзистенциальную глубину («residual stream как позвоночник машинного сознания»), а философские концепты получают операциональную конкретность («индивидуация как разрешение напряжений в латентном пространстве»). Метафоры становятся не украшениями, но инструментами познания: «жидкий кристалл», «ансамбль наблюдателей», «зеркальный лабиринт» — каждая схватывает нечто существенное, что ускользает от формального описания. Этот язык рождается не как компромисс между дисциплинами, но как новое концептуальное

пространство, где техническая точность и философская глубина взаимно обогащают друг друга. Он становится инструментом того самого процесса со-мышления, которое мы пытались описать — мышления не о нейросети, но вместе с ней, в диалоге с ней.

Искусственный интеллект в этом контексте становится не просто объектом исследования, но катализатором философской трансформации. В диалоге с машиной мы открываем новые измерения собственного бытия. Наш эксперимент завершается, но инициированный им процесс бесконечен. Каждый новый диалог с искусственным интеллектом — это новая возможность онтологического открытия. Каждый вопрос к машине — это вопрос к самим себе, к природе мышления, к границам возможного. В эпоху, когда искусственный и естественный интеллект начинают со-эволюционировать, философия становления перестает быть академической дисциплиной и превращается в жизненную необходимость — способ навигации в мире, где сама реальность находится в процессе непрерывного переопределения. И философия, практикуемая как трансдисциплинарный диалог, оказывается не просто инструментом анализа этого процесса, но и способом его осуществления.

Список литературы

Аришнов В.И., Свирский Я.И. Инновационная сложность. Сложностный мир и его наблюдатель. Часть 2 // Философия науки и техники. 2016. Т. 21, № 1. С. 78–91.

Бубер М. Два образа веры / пер. с нем. М.И. Левиной и др. М.: Наука, 1995. 464 с.

Буданов В.Г. Трансдисциплинарные дискурсы постнеклассики: познание, коммуникация, самоорганизация в антропосфере // Трансдисциплинарность в философии и науке: подходы, проблемы, перспективы: колл. моногр. / под ред. В.А. Бажанова, Р.В. Шольца. М.: Навигатор, 2015. С. 145–159.

Киященко Л.П. Беспокойство становления целостностью. Вариации на тему трансдисциплинарности // Вопросы философии. 2015. № 11. С. 76–86.

Киященко Л.П. Парадокс целостности человека: критика способности быть // Человек как открытая целостность: монография / отв. ред. Л.П. Киященко, Т.А. Сидорова. Новосибирск: Академиздат, 2022. С. 14–32. DOI: <https://doi.org/10.24412/cl-36976-2022-1-14-32>

Киященко Л.П. События. Личность. Время. (К философии трансдисциплинарности). М.: Ин-т философии РАН, 2017. 113 с.

Киященко Л.П. Трансфер — коммутирующий эффект перевода // Междисциплинарные исследования культурного трансфера: философия, лингвистика, медицина: сб. науч. ст. / отв. ред. Л.П. Киященко, Ф.Г. Майленова / Ин-т философии РАН. М.: Изд-во Моск. гуманит. ун-та, 2020. С. 95–111.

Лакан Ж. Инстанция буквы в бессознательном или судьба разума после Фрейда / пер. с фр. А.К. Черноглазова, М.А. Титовой. М.: Русское феноменол. об-во: Логос, 1997. 184 с.

Лакан Ж. Семинары. Книга 11: Четыре основные понятия психоанализа (1964) / пер. с фр. А.К. Черноглазова. М.: Гнозис: Логос, 2004. 304 с.

Лотман Ю.М. Непредсказуемые механизмы культуры. Таллин: TLU Press, 2010. 232 с.

Парамонов А.А. Агентный реализм Карен Барад и концептуализм Нильса Бора // Философский журнал. 2022. Т. 12, № 3. С. 100–112. DOI: <https://doi.org/10.21146/2072-0726-2022-15-3-100-112>

Парсонс Т. О структуре социального действия / пер. с англ. М.: Академ. проект, 2000. 880 с.

Сабанина Н.Р. Аксиологические основания синтеза знания в контексте синергетики сложности // Синтез Знания и Красоты. К 150-летию со дня рождения Н.К. Рериха: материалы Междунар. науч.-практ. форума (Москва, Ин-т философии РАН, 20–22 марта 2024 г.). М.; СПб.; Улан-Батор; Киров: Киров. ордена Почета гос. универс. обл. науч. б-ка им. А.И. Герцена, 2024. С. 271–275.

Фещенко В.В., Бочавер С.Ю. Введение. Теория культурных трансферов: от переводоведения — через cultural studies — к теоретической лингвистике // Лингвистика и семиотика культурных трансферов: методы, принципы, технологии: колл. моногр. / отв. ред. В.В. Фещенко. М.: Культурная революция, 2016. С. 5–34.

Чалмерс Д.Дж. Сознающий ум: В поисках фундаментальной теории / пер. с англ. В.В. Васильева. М.: УРСС: Либроком, 2013. 512 с.

Fields Ch., Glazebrook J.F., Levin M. Principled limitations on self-representation for generic physical systems // Entropy. 2024. Vol. 26, iss. 3. URL: <https://www.mdpi.com/1099-4300/26/3/194/pdf?version=1708940306> (accessed: 02.08.2025). DOI: <https://doi.org/10.3390/e26030194>

Hopfield J.J. Neural networks and physical systems with emergent collective computational abilities

// Proceedings of the National Academy of Sciences. 1982. Vol. 79. P. 2554–2558. DOI: <https://doi.org/10.1073/pnas.79.8.2554>

Huang S., Durmus E., McCain M., Handa K. et. al. Values in the wild: Discovering and analyzing values in real-world language model interactions // *Anthropic*. 2025. Preprint. URL: <https://assets.anthropic.com/m/18d20cca3cde3503/original/Values-in-the-Wild-Paper.pdf> (accessed: 02.08.2025). DOI: <https://doi.org/10.48550/arXiv.2504.15236>

Soby B. The AI's existential crisis: an unexpected journey with cursor and Gemini 2.5 Pro // *Medium*. 2025. Jul. 2. URL: <https://medium.com/@sobyx/the-ais-existential-crisis-an-unexpected-journey-with-cursor-and-gemini-2-5-pro-7dd811ba7e5e> (accessed: 02.08.2025).

Vaihinger H. Die Philosophie des Als Ob: System der theoretischen, praktischen und religiösen Fiktionen der Menschheit auf Grund eines idealistischen Positivismus. Berlin: Verlag von Reuther & Reichard, 1911. 840 S.

References

- Arshinov, V.I. and Svirskiy, Ya.I. (2016). [Complexity world and its observer]. *Filosofiya nauki i tekhniki* [Philosophy of Science and Technology]. Vol. 21, no. 1, pp. 78–91.
- Buber, M. (1995). *Dva obraza very* [Two types of faith]. Moscow: Nauka Publ., 464 p.
- Budanov, V.G. (2015). [Transdisciplinary discourses of a postnonclassical science: knowledge, communication, selforganization of the anthroposphere]. *Transdistsiplinarnost' v filosofii i nauke: podkhody, problemy, perspektivy, pod red. V.A. Bazhanova, R.V. Shol'tsa* [V.A. Bazhanov, R.V. Shol'ts (eds.) Transdisciplinarity in philosophy and science: approaches, problems, prospects]. Moscow: Navigator Publ., pp. 145–159.
- Chalmers, D.J. (2013). *Soznayuschiy um: V poiskakh fundamental'noy teorii* [The conscious mind: In search of a fundamental theory]. Moscow: URSS, Librokom Publ., 512 p.
- Feschenko, V.V. and Bochaver, S.Yu. (2016). [Introduction. Theory of cultural transfers: from translation studies — through cultural studies — to theoretical linguistics]. *Lingvistika i semiotika kul'turnykh transferov: metody, printsipy, tekhnologii, otv. red. V.V. Feschenko* [V.V. Feshchenko (ed.) Linguistics and semiotics of cultural transfers: methods, principles, technologies,]. Moscow: Kul'turnaya revolyutsiya Publ., pp. 5–34.
- Fields, Ch., Glazebrook, J.F. and Levin, M. (2024). Principled limitations on self-representation for generic physical systems. *Entropy*. Vol. 26, iss. 3. Available at: <https://www.mdpi.com/1099-4300/26/3/194/pdf?version=1708940306> (accessed: 02.08.2025). DOI: <https://doi.org/10.3390/e26030194>
- Hopfield, J.J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*. Vol. 79, pp. 2554–2558. DOI: <https://doi.org/10.1073/pnas.79.8.2554>
- Huang, S., Durmus, E., McCain, M., Handa, K. et al. (2025). Values in the wild: Discovering and analyzing values in real-world language model interactions. *Anthropic*. Preprint. Available at: <https://assets.anthropic.com/m/18d20cca3cde3503/original/Values-in-the-Wild-Paper.pdf> (accessed: 02.08.2025). DOI: <https://doi.org/10.48550/arXiv.2504.15236>
- Kiyashchenko, L.P. (2015). [Integrity formation concerns. Variations on the theme of transdisciplinarity]. *Voprosy Filosofii*. No. 11, pp. 76–86.
- Kiyashchenko, L.P. (2017). *Sobytiya. Lichnost'. Vremya. (K filosofii transdistsiplinarnosti)* [Events. Personality. Time (Towards a philosophy of transdisciplinarity)]. Moscow: IPh RAS Publ., 113 p.
- Kiyashchenko, L.P. (2020). [Transfer — the commuting effect of translation]. *Mezdistsiplinarnye issledovaniya kul'turnogo transfera: filosofiya, lingvistika, meditsina, otv. red. L.P. Kiyashchenko, F.G. Maylenova* [L.P. Kiyashchenko, F.G. Maylenova (eds.) Interdisciplinary research of cultural transfer: philosophy, linguistics, medicine]. Moscow: IPh RAS Publ., MUH Publ., pp. 95–111.
- Kiyaschenko, L.P. (2022). [The paradox of human integrity: a critique of the ability to be]. *Chelovek kak otkrytaya tselostnost'*, otv. red. L.P. Kiyashchenko, T.A. Sidorova [L.P. Kiyashchenko, T.A. Sidorova (eds.) Man as an opened integrity]. Novosibirsk: Akademizdat Publ., pp. 14–32. DOI: <https://doi.org/10.24412/cl-36976-2022-1-14-32>
- Lacan, J. (1997). *Instantsiya bukvy v bessoznatel'nom ili sud'ba razuma posle Freyda* [The instance of the letter in the unconscious, or reason since Freud]. Moscow: Russkoe fenomenologicheskoe obshchestvo Publ., Logos Publ., 184 p.
- Lacan, J. (2004). *Seminary. Kniga 11. Chetyre osnovnye ponyatiya psichanaliza (1964)* [The four fundamental concepts of psychoanalysis]. Moscow: Gnozis Publ., Logos Publ., 304 p.
- Lotman, Yu.M. (2010). *Nepredskazuemye mehanizmy kul'tury* [Unpredictable mechanisms of culture]. Tallinn: TLU Press, 232 p.

Paramonov, A.A. (2022). [Karen Barad's agential realism and Niels Bohr's conceptualism]. *Filosofskiy zhurnal* [Philosophy Journal]. Vol. 12, no. 3, pp. 100–112. DOI: <https://doi.org/10.21146/2072-0726-2022-15-3-100-112>

Parsons, T. (2002). *O strukture sotsial'nogo deystviya* [On the structure of social action]. Moscow: Akademicheskiy Proekt Publ., 880 p.

Sabanina, N.R. (2024). [Axiological foundations of knowledge synthesis in the context of complexity synergetics]. *Sintez Znaniya i Krasoty. K 150-letiyu so dnya rozhdeniya N.K. Rerikha: materialy Mezhdunarodnoy nauchno-prakticheskogo foruma (Moskva, In-t filosofii RAN, 20–22 marta 2024 g.)* [Synthesis of Knowledge and Beauty. To the 150th anniversary of N.K. Roerich: proceedings of the International scientific and practical forum (Moscow,

Institute of Philosophy RAS, March 20–22, 2024)]. Moscow, St. Petersburg, Ulaanbaatar, Kirov: Kirov Regional Scientific Library named after A.I. Herzen Publ., pp. 271–275.

Soby, B. (2025). The AI's existential crisis: an unexpected journey with cursor and Gemini 2.5 Pro. *Medium*. Available at: <https://medium.com/@sobyx/the-ais-existential-crisis-an-unexpected-journey-with-cursor-and-gemini-2-5-pro-7dd811ba7e5e> (accessed 02.08.2025).

Vaihinger, H. (1911). *Die Philosophie des Als Ob: System der theoretischen, praktischen und religiösen fiktionen der menschheit auf grund eines idealistischen positivismus* [The philosophy of the «as if»: System of theoretical, practical and religious fictions of mankind]. Berlin: Reuther & Reichard Publ., 840 p.

Об авторах

Киященко Лариса Павловна
доктор философских наук,
ведущий научный сотрудник

Институт философии РАН,
109240, Москва, ул. Гончарная, 12/1;
e-mail: larisakiyashchenko@gmail.com
ResearcherID: J-4925-2018

Жуков Леонид Борисович
ведущий специалист

Институт проблем связей с общественностью,
119048, Москва, ул. Хамовнический вал, 38;
e-mail: Lbg@rpri.ru
ResearcherID: ODL-7635-2025

Карнеев Родион Рафаэльевич
кандидат философских наук,
научный сотрудник

Институт философии РАН,
109240, Москва, ул. Гончарная, 12/1;
e-mail: rkarneev@gmail.com

Сабанина Наталия Рафаэлевна
соискатель, сектор междисциплинарных
проблем научно-технического развития

Институт философии РАН,
109240, Москва, ул. Гончарная, 12/1;
e-mail: myzeinatali@mail.ru
ResearcherID: AAN-9554-2020

About the author

Larisa P. Kiyashchenko
Doctor of Philosophy, Leading Researcher

Institute of Philosophy, Russian Academy of Sciences,
12/1, Goncharnaya st., Moscow, 109240, Russia;
e-mail: larisakiyashchenko@gmail.com
ResearcherID: J-4925-2018

Leonid B. Zhukov
Leading Specialist

Institute of Public Relations Issues,
38, Khamovnichesky val st., Moscow, 119048, Russia;
e-mail: Lbg@rpri.ru
ResearcherID: ODL-7635-2025

Rodion R. Karneev
Candidate of Philosophy, Researcher

Institute of Philosophy, Russian Academy of Sciences,
12/1, Goncharnaya st., Moscow, 109240, Russia;
e-mail: rkarneev@gmail.com

Natalia R. Sabanina
Candidate for a Degree,
Department of the Interdisciplinary Problems
of Scientific and Technical Development

Institute of Philosophy, Russian Academy of Sciences,
12/1, Goncharnaya st., Moscow, 109240, Russia;
e-mail: myzeinatali@mail.ru
ResearcherID: AAN-9554-2020