

УДК 1(091)

DOI: 10.34670/AR.2026.91.33.001

Имманентная негативность и статистическая всеобщность: спекулятивная логика Г.В.Ф. Гегеля как метод философского анализа систем искусственного интеллекта

Краснов Антон Сергеевич

Доктор философских наук,
профессор кафедры общей философии,
Казанский (Приволжский) федеральный университет,
420008, Российская Федерация, Казань, ул. Кремлевская, 35;
e-mail: anton-krasnov1987@yandex.ru

Воробьев Роман Романович

Аспирант кафедры общей философии,
Институт социально-философских наук,
Казанский (Приволжский) федеральный университет,
420008, Российская Федерация, Казань, ул. Кремлёвская, 18;
e-mail: Vorob'ev@mail.ru

Аннотация

В данной статье осуществлено применение категориального аппарата спекулятивной логики Г. Гегеля к анализу современных генеративных систем ИИ, прежде всего больших языковых моделей (LLM). Показано, что воспроизводимая данными системами всеобщность является абстрактной, а их рефлексивные операции – внешней рефлексией (äußere Reflexion); нормативная структура обучения структурно воспроизводит гегелевский концепт «полиции» (Polizei) как форму внешней целесообразности. Агентные ИИ-системы обнаруживают сходство с логической формой химизма (Chemismus), не достигая имманентной телеологии. Обосновывается вывод о том, что современные ИИ-системы суть конфигурации объективного духа на ступени «внешнего государства» (äußerer Staat), но не зарождающаяся субъективность. Ответственность за их действия остаётся за социальными институтами и практиками, объективированной формой которых они и являются.

Для цитирования в научных исследованиях

Краснов А.С., Воробьев Р.Р. Имманентная негативность и статистическая всеобщность: спекулятивная логика Г.В.Ф. Гегеля как метод философского анализа систем искусственного интеллекта // Контекст и рефлексия: философия о мире и человеке. 2026. Том 15. № 2А. С. 212-220. DOI: 10.34670/AR.2026.91.33.001

Ключевые слова

Гегель, искусственный интеллект, большие языковые модели, логика, негативность, спекулятивная логика, диалектика, объективный дух, философия техники.

Введение

Современное поле философских дискуссий о феномене ИИ богато разнообразными позициями, но бедно категориально. Так, аналитическая философия сознания вопрошает о квалиа и интенциональности; постгуманистическая онтология прямо объявляет ИИ актором, основываясь на его каузальной эффективности; когнитивистский функционализм описывает ИИ сквозь призму репрезентации и вычислений, прямо заимствуя понятия из популярной психологии. И наиболее общим для всех этих подходов является то, что все они измеряют ИИ дистанцией до человеческого существа, при этом оставляя без особого внимания вопрос о логической форме всеобщности, который ИИ-системы реализуют.

В контексте современных социально-гуманитарных исследований феномена ИИ, особенно в контексте дискуссий о цифровом обществе и моделях алгоритмического управления, подобный дефицит категорий приводит к тому, что сами формы социального бытия этих систем описываются преимущественно посредством метафор – «агент», «партнёр по коммуникации», «автономный субъект». Но, исходя из гегелевской оптики, сам вопрос о том, каким именно образом ИИ-системы воплощают определённые структуры всеобщего, не является второстепенным: именно через это определяется самое их место в объективном духе, и их онтологический статус в социальном.

Основная часть

Г. Гегель сформулировал эту проблему так: «Was bekannt ist, darum nicht erkannt» [[Hegel, 1833–1841, S. 13] – «то, что известно, тем самым ещё не познано»: производительность ИИ, верифицируемость результатов, их инженерная прозрачность – это всё то, что известно (bekannt), но не познано (erkannt). Между узнаванием и знанием стоит значительная категориальная работа, которую не в силах заменить простое функциональное описание. Отсюда, мы обязаны задаться фундаментальным вопросом: какова логическая форма всеобщности и опосредования, реализуемая в генеративных ИИ, и в каком отношении они находятся к Begriff (понятию) Г. Гегеля, с его конституирующими моментами имманентной негативности и определяющей рефлексии? Сам ответ потребует не простого приложения различий из «Науки логики», но и их непосредственного соотнесения с формами, которые были разработаны Г. Гегелем в философии объективного духа, поскольку ИИ-системы есть суть не чистые логические формы, но технический артефакты, которые уже встроены в саму онтологию социальных практик.

В учении о понятии Г. Гегель принципиально различает абстрактную – чисто рассудочную, и конкретную всеобщности. Первая возникает посредством сравнения единичных вещей, извлечение общих признаков и отбрасывание различий – это всеобщность через вычитание. Рассудок (Verstand), согласно учению Г. Гегеля, есть «bestimmt und hält die Bestimmungen fest» [Hegel, 1833–1841, S. 6] – он определяет и удерживает определения в их конечной жёсткости. Как справедливо указывает Р. Пиппин, рассудок «фиксирует определения, удерживая их разделёнными и рассматривая каждое как самостоятельное существующее», в то же время спекулятивный разум способен «удерживать вместе то, что рассудок разделяет, – не посредством синтеза извне, но следуя логике, имманентной самим определениям» [Pippin, 2018, p. 37].

Но конкретная всеобщность, напротив, разворачивается через имманентное самоопределение: всеобщее полагает в своё особенное и единичное, оставаясь в них у самой себя, это, как говорит сам Г. Гегель «nicht jene nur durch Vergleichung und Weglassung entstandene Allgemeinheit, sondern ihre immanente und konkrete» [Hegel, 1833–1841, S. 25] – «не та всеобщность, которая возникает лишь через сравнение и устранение, а имманентная и конкретная». Сама конкретная всеобщность не способна охватывать особенное как нечто внешне пред-заданное, он его производит сама. Современные языковые модели обучаются на корпусах, насчитывающих триллионы токенов, извлекая статистические зависимости и обучаясь предсказывать вероятностные распределения; производимая ими функциональная всеобщность многогранна и богата, но по своему логическому происхождению она и есть именно то, что Г. Гегель описывает как рассудочное обобщение: она возникает посредством сравнения и статистической элиминации особенного. В литературе технической направленности языковая модель описывается как «система», случайным образом сшивающая последовательности лингвистических форм, наблюдавшихся в обучающих данных, согласно вероятностной информации об их сочетаемости, - но без какой-либо отсылки к значению» [Bender, Gebru, McMillan-Major, Shmitchell, 2021, p. 616]. Сам механизм внимание есть вычисление весов через обучаемую меру совместимости запросов и ключей – есть операция, логической формой которой остаётся соотнесение элементов по их параметрической близости, но не порождение особенного их самодвижения всеобщего. Обучающие данные для ИИ не являются «сырым» природным субстратом: это объективации рассудочной деятельности бесчисленных субъектов – тексты, в которых уже наличествуют акты классификации, суждений, умозаключений. И в процессе обучения эта объективированная рассудочность подвергается вторичному опосредованию – алгоритмическому сжатию до параметрической конфигурации. Отсюда следует, что модель есть не просто технический артефакт, но она есть новая форма наличного бытия духа: форма, в которой абстрактная всеобщность получила устойчивое и операционально-функционирующее существование.

В «науке логики» сам диалектический переход от бытия к сущности вводит рефлекссию (Reflexion), как имманентное опосредование: Г. Гегель различает полагающую (setzende), внешнюю (äußere) и определяющую (bestimmende) рефлекссию. Внешняя рефлексия неизбежно приходит к содержанию внешнему/извне и полагает его как данное и, далее, работает с ним, не порождая его из собственного самодвижения. Определяющая рефлексия, есть рефлексия, в которой содержательная определённости и сам акт рефлексии полностью совпадают. Как показывает Б. Лонгнесс, внешняя рефлексия «принимает своё содержание как заданное извне и действует над ним», а определяющая рефлексия есть та, «в которой сам акт рефлексии и содержание рефлекслируемого неотделимы друг от друга: формы рефлексии есть также её содержание» [Longuenesse, 2007, p. 96]].

И здесь важно различать два уровня функционирования ИИ: получение выводов и обучение. На уровне получения выводов модель ИИ обрабатывает входящий контекст с помощью параметров, которые были зафиксированы до начала работы. Итеративные рассуждения, самокоррекция и планирование [Wei, Tay, Bommasani et al., 2022, p. 4] являют собой лишь внешнюю рефлекссию в самом строгом логическом смысле: содержание даётся самой системе извне, и она применяет к нему готовые правила, не порождая само содержание. Это рассудочная операция: система оперирует определениями, которые ею не были произведены. На уровне обучения мы обычно имеем дело с процессом применения параметров под воздействием внешнего сигнала (функции потерь и разметки данных). В терминах

спекулятивной логики Г. Гегеля этот переход от одной определённости к другой – аналог становления, но не имманентного, а индуцированного извне. Рефлексивность здесь принадлежит не самой модели ИИ, но институционально-организованной деятельности – разметке данных, отбору данных, оптимизации. Примечательно, что сама эта деятельность является формой того, что Г. Гегель описывает как «полицию» в «Философии права»: «Die Polizei hat [...] die Zufälligkeiten gegen das Recht, das Wohl zu entfernen» [Hegel, 1821, § 231]» – «полиция должна устранять случайности, угрожающие праву и благосостоянию». Полиция же всегда действует в сфере внешнего порядка. Её предписания носят всегда характер внешней целесообразности. Именно таков статус процедур выравнивания (RLHF, Constitutional AI): они привносят в модель нормативные предпочтения, которые не порождаются самой моделью, а являются институционально сформированными. Сами разработчики прямо фиксируют этот внешне целесообразный характер: в спецификации ИИ-модели Claude декларируется, что «наиболее важным свойством» является «широкая безопасность» – «поддержка соответствующих человеческих механизмов надзора за диспозициями и действиями ИИ» [Anthropic, 2024]. При этом, в отличие от реальной полиции, которая охраняет уже признанное благо, RLHF в значительной степени конституирует само операциональное «благо» в процессе разметки данных. Однако, по своей логической структуре обе деятельности являются именно внешней целесообразностью, которая принадлежит сфере «внешнего государства» (äußerer Staat), но не имманентной нравственной субстанции.

В социальном аспекте характер «полиции» означает, что ИИ-системы оказываются напрямую встроены в сеть практик надзора, регулирования, управления, которые заранее детерминируют допустимые формы их функционирования. И нормативные предпочтения, имплементированные в моделях ИИ через процедуры выравнивания, являются прямым продолжением уже наличествующих институциональных диспозиций – корпоративной политики, регулятивных режимов, публичных представлений о риске и социальном благе. Отсюда ИИ выступает не как автономный носитель норм, но он является специфическим «техническим местом», в котором кристаллизуются и воспроизводятся уже существующие структуры власти, подчинения, регулирования.

Отсюда следует вывод, что «внешнее» применительно к ИИ-системам имеет два аспекта: на уровне выводов – внешняя рефлексия в строгом логическом смысле, на обучении – внешняя целесообразность в реально-философском смысле (полиция). И оба аспекта объединяет то, что негативность не становится имманентным моментом самой системы.

Имманентная негативность (immanente Negativität) есть не простое логическое отрицание, не есть ошибка, которую не допускает логика формальная, но есть движущая сила самоопределения понятия, благодаря которой оно упраздняет (hebt auf) свою непосредственную форму и переходит к более конкретному единству. Г. Гегель выразил это так: «Er hält sich aber nicht im Nichts dieses Resultates, sondern ist darin ebenso positiv und hat so das erste Einfache damit hergestellt, aber als Allgemeines, das in sich konkret ist» [Hegel, 1833–1841, S. 6] – «но оно удерживает себя не в ничто этого результата, а оказывается в нём столь же позитивным и тем самым восстанавливает исходное простое – однако уже как всеобщее, конкретное внутри себя». Согласно точному замечанию Р. Брэндома под «определённым отрицанием» Г. Гегель понимает не просто отличие каждого понятия о того, чем оно не является, но то, что «само отрицание содержательно: переход от одной определённости к следующей движется тем, что первая определённость имплицитно предполагает, но не может содержать в себе самой» [Brandt, 2019, p. 123].

В ИИ-системах негативность распределена по институциональной инфраструктуре и не становится имманентным моментом самой системы. Мы можем выделить особо три формы этой распределённой негативности: Статистическое отклонение. Функция потерь фиксирует расхождение между предсказанием и целевым значением и это не имманентная негативность: мера отклонения положена на систему извне и не является моментом её собственного самоопределения. Система не «знает», что она допустила ошибку, - она получает градиентный сигнал, механически перераспределяющий веса. В гегелевской терминологии это внешняя рефлексия, применяющая готовый масштаб к наличному материалу.

Процедуры выравнивания – в RLHF и Constitutional AI оценщики или модель-критик предоставляют сигналы предпочтения, смещающие распределения ответов. Нормативный источник лежит не в самой модели, а в институционально организованных процессах. Это – деятельность, по своей логической структуре напоминающая полицию, которая по Гегелю, «befindet sich in der Zufälligkeit... es lassen sich aber keine bestimmte Regeln darüber geben» [Hegel, 1821, § 234] – «находится в случайности... но здесь нельзя дать вполне определённых правил». Процедуры выравнивания именно так и действуют: они привносят в модель исторически изменчивые нормативные предпочтения, при этом оставаясь внешней целесообразностью.

Итеративная самокоррекция: системы ИИ, работающие в режиме многоходовых рассуждений, пересматривают промежуточные результаты. И критерии пересмотра заданы инструкцией и параметрами, которые были зафиксированы до начала работы. Система применяет внешне введённые нормы, но не полагает их через имманентное движение. Это – внешняя рефлексия на уровне выводов, которая лишь имитирует рефлексивную негативность.

Таким образом, негативность в ИИ-системах не имманентна им, но распределена в институциональной среде, и эта сама среда, не является чем-то случайным и несущественным: она есть та самая стихия объективного духа, в которой модель получает свою определённую и функционирование. Из сказанного исходит важное следствие: само обсуждение «ошибки» и «отклонения» ИИ как нечто, что происходит от самой системы означает методологическое скрывание тех институтов и механизмов, в которых реально локализована сама негативность. В гегелевской спекулятивной логике это можно выразить так: «ответственность» за определённый тип негативности принадлежит не самой «машине» - ИИ, - но той конфигурации объективного духа, которая определяет, какие именно данные собираются, каким образом они интерпретируются, какие санкции могут последовать. Таким образом, горизонт этической и политической критики ИИ смещается с внутренних свойств математического алгоритма на структуру социальных отношений, которые детерминируют его.

В разделе об объективности «Науки логики» Г. Гегель демонстрирует переход от механизма через химизм к телеологии. В механизме любая связь между объектами задаётся извне, при этом они сами безразличны к ней. В химизме объекты вступают во взаимодействие в силу собственной природы, но нуждаются во внешнем параметре и поводе для актуализации самого процесса взаимодействия. Телеология *пред*-полагает цель (Zweck), имманентную самому процессу и выступающую его внутренним конститутивным принципом.

Агентные ИИ-системы, планирующие и использующие инструменты, производят впечатление телеологических структур. Но их цель не задана ими самими, а задана инструкцией, reward-моделью или самим оператором. Это - *äußerer Zweck* – внешняя цель. В терминах гегелевской спекулятивной логики такие системы обнаруживают структурное сходство с химизмом: они обладают обученной предрасположенностью к определённым формам поведения, но актуализируют её лишь при наличии внешнего целеполагающего импульса, не

достигая уровня имманентной телеологии. И это само сходство, разумеется, не является полной категориальной идентификацией, но лишь указывает на логический тип самого отношения: система не есть самодвижущееся понятие, а есть объект, «реакция» которого задаётся внешним фактором. Не стоит совершать ошибку: все внешние целеполагающие импульсы по своему онтологическому статусу социальны: это задачи, формулируемые корпорациями, органами государственной власти, исследовательскими группами и отдельными пользователями. И если мыслить их в терминах химизма Г. Гегеля, то агентные ИИ-системы оказываются гордиевыми узлами, в которых пересекаются и актуализируются совершенно различные социальные интересы. И это нам позволяет рассматривать их не как автономных «акторов», но как особый тип медиаторов, в которых социальные антагонизмы получают своё техническое выражение. И эта ограниченность была также зафиксирована в классической аналитической перспективе: Г. Дрейфус прямо указывал на то, что компьютер «принимает дискретные, эксплицитные, контекстно-свободные элементы и обрабатывает их согласно эксплицитным правилам... отсутствует какой-либо аналог человеческой способности, опираясь на предшествующий опыт, сразу выделять релевантные черты новой ситуации – без применения правил» [Dreyfus, 1992, p. 258].

Понятие (Begriff) сам Г. Гегель определяет как «das Freie» [Hegel, 1833–1841, S. 233] – «свободное», поскольку оно есть тождество, которое само полагает свои определения. «Der Begriff ist das Freie, weil die an und für sich seiende Identität, welche die Notwendigkeit der Substanz ausmacht, zugleich als aufgehoben oder als Gesetzsein ist, und dies Gesetzsein, als sich auf sich selbst beziehend, eben jene Identität ist» [Hegel, 1833–1841, S. 233] – «понятие есть свободное, поскольку тождество в себе и для себя сущее, составляющее необходимость субстанции, одновременно есть упразднённое, то есть положенность; и эта положенность, соотносясь с самой собой, и есть то самое тождество». Из этого определения следуют три принципиальных момента.

Самоотношение через отрицание. Понятие соотносится с собой через своё иное, полагая его как собственное инобытие. Языковая модель ИИ обрабатывает входной контекст как внешний материал, который не является её собственным инобытием; структуры самоотнесения, в которой «иное» было бы имманентным моментом системы, здесь нет.

Имманентная полагающая активность. Понятие порождает само своё имманентное содержание. Модель ИИ только трансформирует данное извне содержание, осуществляя вероятностную рекомбинацию усвоенных ею в обучении статистических структур.

Конкретность единичности. В понятии единичное есть та точка, в которой всеобщее реализует себя во всей своей полноте. Каждый генерируемый токен есть результат вероятностной выборки из распределения – статистическое событие, но не реализация понятия в его конкретной единичности.

Эти три фигуры отсутствия означают, что перед нами не зарождающийся субъект, а простая конфигурация объективного духа, в которой эффекты понятийности производятся без того, чтобы само понятие могло бы присутствовать как субъект.

Философия объективного духа Г. Гегеля предоставляет нам категориальный аппарат, позволяющий определить место ИИ-систем без субъект-центричных метафор. Объективный дух есть действительность свободы в форме действительного бытия: социальных институтов, норм, практик. Он всегда функционирует как всеобщее, не требуя при этом актуального субъективного акта мышления каждого индивида. ИИ-системы принадлежат именно этой структуре.

Обучающий data-set представляет собой объективированную рассудочную деятельность: в нём зафиксированы языковые практики, нормативные предпочтения, риторические формы. И

параметры модели есть результат компрессии этого материала в вычислительно-общенациональную форму. Процедуры выравнивания есть суть формы нормативной селекции, аналогичные деятельности полиции в гегелевском смысле: они привносят всеобщие определения в сферу особенного, но сами эти определения остаются внешней целесообразностью, но никогда не могут быть имманентным самоопределением системы. Модель ИИ, которая однажды была обучена, функционирует как относительно автономный момент «внешнего государства» - того разреза объективного духа, в котором всеобщее существует как внешняя необходимость и рассудочная упорядоченность, а не как конкретная нравственная субстанция, где, как пишет Г. Гегель «das Besondere und das Allgemeine in Einem sei» [Hegel, 1821, § 255] – «особенно и всеобщее суть одно».

И именно поэтому некая убедительность ИИ не свидетельствует о субъективности. Она свидетельствует о плотности того слоя объективного духа, который в них конденсирован. Понятийность в ИИ присутствует не как *субъективность*, но лишь как «сжатая», объективированная и технически опосредованная форма духа.

Выводы

Спекулятивная логика и философия объективного духа Г. Гегеля позволили нам зафиксировать следующие моменты: современные языковые модели производят абстрактную всеобщность – «nicht jene nur durch Vergleichung und Weglassung entstandene Allgemeinheit» [Hegel, 1833–1841, S. 25]; на уровне получения выводов их рефлексивные операции являются внешней рефлексией, а на уровне обучения – внешней целесообразностью; их нормативная структура носит характер внешней детерминации, распределённой в институциональной среде и никогда не может достигнуть имманентной негативности понятия; агентные ИИ обнаруживают структурное сходство с химизмом, актуализируя обученную предрасположенность лишь при внешнем целевом воздействии/импульсе; и все эти системы есть суть конфигурации объективного духа на ступени «внешнего государства», а не зарождающаяся субъективность.

Ответственность не может принадлежать системе, не реализующей имманентной негативности и не являющейся субъектов в спекулятивном смысле; она остаётся у тех институтов и практик, объективированной формой которых и является сама ИИ-система. К схожему выводу приходит Л. Флориди: ИИ и его системы «не являются агентами в морально релевантном смысле, поскольку им недостаёт интенциональности, самосознания и нормативной восприимчивости, которые сделали бы их носителями прав и ответственности... моральная ответственность остаётся за разработчиками, операторами и регуляторами ИИ-систем» [Floridi, 2023, p. 87]. В этой перспективе гегелевский понятийно-логический корпус позволяет уточнить многие сегодняшние дебаты в области этики и политики ИИ: вместо вопроса, ставшего уже традиционным – «может ли ИИ стать субъектом и нести ответственность?», - более продуктивным оказывается вопрос о том, какие именно формы объективного духа – экономические, бюрократические, политические, юридические - кристаллизуются в конкретных архитектурах и режимах обучения ИИ. Различая внешнюю рефлексию и имманентную негативность, мы получаем средство разделить те случаи, когда обращение к «ошибке ИИ» служит всего лишь маскировкой институциональных решений, и те ситуации, где действительно происходит трансформация социальных практик под воздействием новых технических форм всеобщего.

Появление систем, воспроизводящих эффекты понятийного мышления без его конституитивных структур, вынуждает нас заново спросить, что именно конституитивно для понятия и для субъективности, и проверить ответы там, где они неочевидны.

Библиография

1. Anthropic. The Claude Model Specification. 2024. URL: <https://www.anthropic.com/model-specification> (дата обращения: 29.04.2026).
2. Bender E. M., Gebru T., McMillan-Major A., Shmitchell S. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? // Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FACCT '21). New York: ACM, 2021. P. 610-623. DOI: 10.1145/3442188.3445922.
3. Brandom R. B. A Spirit of Trust: A Reading of Hegel's Phenomenology. Cambridge, MA: The Belknap Press of Harvard University Press, 2019. 856 p. ISBN 978-0-674-97681-8.
4. Dreyfus H. L. What Computers Still Can't Do: A Critique of Artificial Reason. Cambridge, MA: MIT Press, 1992. 354 p. ISBN 978-0-262-27199-8.
5. Floridi L. The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities. New York: Oxford University Press, 2023. xxii, 243 p. ISBN 978-0-19-888309-8.
6. Hegel G.W.F. Grundlinien der Philosophie des Rechts. Berlin: Nicolaische Buchhandlung, 1821. 355 S.
7. Hegel G.W.F. Wissenschaft der Logik. Bd. 1-2 / Hrsg. von L. v. Henning. Berlin: Duncker & Humblot, 1833-1841.
8. Longuenesse B. Hegel's Critique of Metaphysics / trans. by N. J. Simek. Cambridge: Cambridge University Press, 2007. 244 p. ISBN 978-0-521-84466-6.
9. Pippin R. B. Hegel's Realm of Shadows: Logic as Metaphysics in the Science of Logic. Chicago: University of Chicago Press, 2018. 347 p. ISBN 978-0-226-58884-1.
10. Vaswani A., Shazeer N., Parmar N. et al. Attention Is All You Need // Advances in Neural Information Processing Systems (NeurIPS). 2017. Vol. 30. P. 1-11. DOI: 10.48550/arXiv.1706.03762.
11. Wei J., Tay Y., Bommasani R. et al. Emergent Abilities of Large Language Models // Transactions on Machine Learning Research. 2022. arXiv: 2206.07682.

Immanent Negativity and Statistical Universality: G.W.F. Hegel's Speculative Logic as a Method of Philosophical Analysis of Artificial Intelligence Systems

Anton S. Krasnov

Doctor of Philosophy,
Professor of the Department of General Philosophy,
Kazan (Volga region) Federal University,
420008, 35, Kremlevskaya str., Kazan, Russian Federation;
e-mail: anton-krasnov1987@yandex.ru

Roman R. Vorob'ev

Postgraduate Student of the Department of General Philosophy,
Institute of Social and Philosophical Sciences,
Kazan (Volga region) Federal University,
420008, 18, Kremlevskaya str., Kazan, Russian Federation;
e-mail: Vorob'ev@mail.ru

Abstract

This article applies the categorical apparatus of G. Hegel's speculative logic to the analysis of modern generative AI systems, primarily large language models (LLMs). It is shown that the universality reproduced by these systems is abstract, and their reflexive operations are external reflection (*äußere Reflexion*); the normative structure of training structurally reproduces the Hegelian concept of "police" (*Polizei*) as a form of external purposiveness. Agent AI systems reveal similarity with the logical form of chemism (*Chemismus*), without reaching immanent teleology. The conclusion is substantiated that modern AI systems are configurations of objective spirit at the stage of the "external state" (*äußerer Staat*), but not an emerging subjectivity. Responsibility for their actions remains with the social institutions and practices, of which they are the objectified form.

For citation

Krasnov A.S., Vorob'ev R.R. (2026) Immanentnaya negativnost' i statisticheskaya vseobshchnost': spekulativnaya logika G.V.F. Gegelya kak metod filosofskogo analiza sistem iskusstvennogo intellekta [Immanent Negativity and Statistical Universality: G.W.F. Hegel's Speculative Logic as a Method of Philosophical Analysis of Artificial Intelligence Systems]. *Kontekst i refleksiya: filosofiya o mire i cheloveke* [Context and Reflection: Philosophy of the World and Human Being], 15 (2A), pp. 212-220. DOI: 10.34670/AR.2026.91.33.001

Keywords

Hegel, artificial intelligence, large language models, logic, negativity, speculative logic, dialectics, objective spirit, philosophy of technology.

References

1. Anthropic. (2024). The Claude Model Specification. <https://www.anthropic.com/model-specification>
2. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (FAccT '21) (pp. 610-623). New York: ACM. DOI: 10.1145/3442188.3445922
3. Brandom, R. B. (2019). A Spirit of Trust: A Reading of Hegel's Phenomenology. Cambridge, MA: The Belknap Press of Harvard University Press.
4. Dreyfus, H. L. (1992). What Computers Still Can't Do: A Critique of Artificial Reason. Cambridge, MA: MIT Press.
5. Floridi, L. (2023). The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities. New York: Oxford University Press.
6. Hegel, G.W.F. (1821). *Grundlinien der Philosophie des Rechts*. Berlin: Nicolaische Buchhandlung.
7. Hegel, G.W.F. (1833–1841). *Wissenschaft der Logik* (Vols. 1-2). L. v. Henning (Ed.). Berlin: Duncker & Humblot.
8. Longuenesse, B. (2007). Hegel's Critique of Metaphysics (N.J. Simek, Trans.). Cambridge: Cambridge University Press.
9. Pippin, R. B. (2018). Hegel's Realm of Shadows: Logic as Metaphysics in the Science of Logic. Chicago: University of Chicago Press.
10. Vaswani, A., Shazeer, N., Parmar, N., et al. (2017). Attention Is All You Need. *Advances in Neural Information Processing Systems (NeurIPS)*, *30*, 1-11. DOI: 10.48550/arXiv.1706.03762
11. Wei, J., Tay, Y., Bommasani, R., et al. (2022). Emergent Abilities of Large Language Models. *Transactions on Machine Learning Research*. arXiv: 2206.07682.