

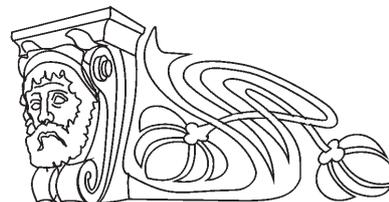


Известия Саратовского университета. Новая серия. Серия: Философия. Психология. Педагогика. 2023. Т. 23, вып. 4. С. 414–418
Izvestiya of Saratov University. Philosophy. Psychology. Pedagogy, 2023, vol. 23, iss. 4, pp. 414–418
<https://phpp.sgu.ru> <https://doi.org/10.18500/1819-7671-2023-23-4-414-418>, EDN: PBDAAR

Научная статья
УДК 004.8+[1:316]

Социально-философские аспекты развития генеративного искусственного интеллекта

М. А. Шаткин



Саратовский национальный исследовательский государственный университет имени Н. Г. Чернышевского, Россия, 410012, г. Саратов, ул. Астраханская, д. 83

Шаткин Максим Александрович, кандидат философских наук, доцент кафедры философии и методологии науки, maximshatkin@gmail.com, <https://orcid.org/0000-0002-8138-6819>

Аннотация. Введение. Статья фокусируется на осмыслении развития генеративного искусственного интеллекта (ГИИ) и динамики его взаимодействия с обществом с точки зрения социальной философии с помощью диалектического метода. **Теоретический анализ.** Особенностью ГИИ является обусловленная его сложностью непостижимость для человеческого сознания. ГИИ, который генерирует контент без его понимания, отражает особенность человеческого творчества, создающего феномены без полного понимания их принципов и последствий. Непредсказуемый ГИИ, чьи отдельные возможности существенно превосходят человеческие, воспринимается не как инструмент, а как помощник. Однако универсальный характер ГИИ препятствует формированию личной истории его взаимодействия с пользователями, что приводит к потребности к обеспечению эксклюзивности ГИИ через связь с материальным носителем. Благодаря этому ГИИ обретает опыт самостоятельного наличного бытия, запускающий цикл бесконечного самосовершенствования и самопорождения, а также трансформацию материи, включая общество. Результатом такой трансформации может стать цифроматериальное общество роевого интеллекта, где взаимная открытость сознаний людей поддерживается их связью с ГИИ. **Заключение.** Развитие ГИИ представляет диалектическое движение от абстрактной универсальности в сторону конкретности, материального воплощения, и затем обратно к конкретной универсальности через включение и трансформацию в рамках ГИИ материального мира.

Ключевые слова: генеративный искусственный интеллект, нейронные сети, общество, умная материя, роевый интеллект

Для цитирования: Шаткин М. А. Социально-философские аспекты развития генеративного искусственного интеллекта // Известия Саратовского университета. Новая серия. Серия: Философия. Психология. Педагогика. 2023. Т. 23, вып. 4. С. 414–418. <https://doi.org/10.18500/1819-7671-2023-23-4-414-418>, EDN: PBDAAR

Статья опубликована на условиях лицензии Creative Commons Attribution 4.0 International (CC-BY 4.0)

Article

Socio-philosophical aspects of the development of generative artificial intelligence

M. A. Shatkin

Saratov State University, 83 Astrakhanskaya St., Saratov 410012, Russia

Max Shatkin, maximshatkin@gmail.com, <https://orcid.org/0000-0002-8138-6819>

Abstract. Introduction. The article offers a socio-philosophical reflection on the development of generative artificial intelligence (GAIN) and its interaction with human society, through the lens of the dialectical method. **Theoretical analysis.** The characteristic feature of the GAIN is its complexity, which makes it incomprehensible to human consciousness, making it similar to social institutions created by society but not controlled by it. The GAIN, which creates content without understanding it, reflects the characteristic of the human spirit, which creates phenomena without ever understanding the principles and consequences of their work. The possibilities and unpredictability of the GAIN's work, which exceeds the capabilities of the human mind, lead to the emergence of an attitude towards it not as a tool, but as a helper. But the universal nature of the GAIN prevents the generation of a personal history of interaction between a person and the GAIN, which leads to the need for individualization and concretization of the GAIN in a material medium. The connection of GAIN with a certain material medium gives GAIN the experience of independent existence, which triggers the cycle of infinite self-perfection and self-generation, as well as the transformation of matter, including society. The result of such transformation can be a digital-material society of swarm intelligence, where the openness of people's consciousnesses to each other is conditioned by their connection with GAIN. **Conclusion.** The development of the GAIN is a dialectical movement from universality towards its concreteness and material embodiment, and then back to universality through the inclusion and transformation of the material world within the GAIN.

Keywords: generative artificial intelligence, neural networks, society, intelligent matter, swarm intelligence

For citation: Shatkin M. A. Socio-philosophical aspects of the development of generative artificial intelligence. *Izvestiya of Saratov University. Philosophy. Psychology. Pedagogy*, 2023, vol. 23, iss. 4, pp. 414–418 (in Russian). <https://doi.org/10.18500/1819-7671-2023-23-4-414-418>, EDN: PBDAAR
This is an open access article distributed under the terms of Creative Commons Attribution 4.0 International License (CC-BY 4.0)



Введение

Изобретение генеративного искусственного интеллекта (ГИИ) – нейронных сетей, создающих по запросу человека новый мультимедийный или текстовый контент, – часто сравнивается с изобретением электричества, книгопечатания, алфавита или колеса. Однако наиболее близким аналогом этого изобретения является приручение огня [1, р. 6–7], поскольку в ГИИ была «приручена» такая «природная сила», как вероятность и ее подсчет. В ГИИ человек взаимодействует с законами вероятности, которые оперируют не физическими объектами, а символическими элементами человеческой культуры, структурируемыми не посредством Логоса, а через безличный подсчет вероятностей. Так же как огонь часто выходит из-под контроля человека, так и этот подсчет часто дает результат за пределами человеческих ожиданий. Для человечества, которое является статистической аномалией в обозримой Вселенной и при этом активно увеличивает вероятность собственного исчезновения, это несет риск того, что в один момент сверхмощный ГИИ может прийти к выводу о том, что человечество является статистической погрешностью и его интересы не нужно учитывать.

Подобные алармистские рассуждения отражают линейный, не-диалектический взгляд на становление ГИИ. Актуальна задача социально-философского осмысления взаимодействия ГИИ и человека (как отдельного индивида, так и общества) в его диалектическом развитии. С учетом того, что в лице ГИИ мы имеем дело с нечеловеческим и даже сверхчеловеческим интеллектом, целесообразно обращение к религиозно-философскому дискурсу, обладающему богатым опытом описания подобных взаимодействий. В данной статье предлагается социально-философское осмысление перспектив развития ГИИ с точки зрения его взаимодействия с человеческим обществом, с использованием элементов религиозно-философского дискурса и с опорой на диалектический метод.

Теоретическая анализ

1. *Непонимающее творение.* ГИИ функционирует в форме нейронных сетей, генерирующих по запросу пользователей ответы на поставленные вопросы, а также мультимедийный контент. С этой точки зрения ГИИ представляет собой результат человеческого творчества в нескольких смыслах. ГИИ появился как результат обучения

нейронных сетей на огромных массивах данных, созданных человечеством (примечательно, что данные естественных и точных наук могут обеспечить только создание алгоритмов, но не самообучающихся нейронных сетей). Поэтому «создание» ГИИ человеком не является «созданием», подобно созданию компьютера или ткацкого станка: ГИИ по специфике происхождения ближе к социальным институтам (государство, образование, «невидимая рука рынка» и т.д.), чья склонность превращаться в «джинна, выпущенного из бутылки», рефлексировалась еще в XVIII в. [2, с. 13–14]. ГИИ возник как отражение человеческого творчества в цифровой среде, где «отражение» означает повторение (ГИИ опирается на накопленные человечеством знания и накопленные Интернетом высказывания людей, в том числе не имеющие смысла), обобщение (в своих ответах ГИИ не повторяет отдельные высказывания, на которых он обучался, а формирует собственные, которые он счел наиболее вероятными, и сублимацию (несмотря на свою осмысленность, создаваемые ГИИ тексты и контент не отражают конкретную материальную реальность или жизненную ситуацию, а являются наиболее оптимальным вариантом).

Сложность ГИИ делает его недоступным для понимания человека аналогично тому, как в прежние эпохи социальные и культурные институты были непостижимы для человека в качестве его творения, а сегодня большая часть населения пользуется современными технологиями без понимания всех принципов их работы. Для религиозно-философского дискурса разрешением этого противоречия служила концепция со-творчества человека Богу, т.е. задачи человека продолжать Божественное творение в мире. Эта концепция предполагает образ человека, который не понимает всего смысла своего творчества, поскольку оно является частью Божественного замысла о творении, а он не может быть постигнут тварными существами. В этом смысле ГИИ воплощает такую черту человеческого творчества, как *непонимающее творение*: создание нового без четкого понимания того, как оно возникает и работает, к каким новым (эмерджентным) явлениям приведет, какие возможности даст и как обратно повлияет на человека. Применительно к ГИИ непонимающее творение имеет два аспекта: а) человек не понимает, как работает ГИИ; б) ГИИ создает новое, не имея на данный день возможности его понять.

2. *Общество и сверхъестественная сущность.* Знание человека о ГИИ уже сегодня ограничивается знанием об абстрактных ха-



рактических характеристиках нейронных сетей (количество параметров). Это знание порождает веру в то, что ГИИ обладает силой сделать то (или даже «все»), что рядовому пользователю затруднительно или невозможно, и прежде всего это достижение эффективности – в создании и обработке контента, логистике, управлении сложными системами и ведении войны. Непонимание того, как именно ГИИ приходит к своим ответам, превышающим степень разумности среднего пользователя, делает ГИИ «для-нас» одной из форм сверхъестественных сущностей, обращение к способностям которых помогает человеку и обществу в решении их проблем. Потенциальная возможность взлома вражеского ГИИ и его использование против противника для общества может стать возрождением древнеримских ритуалов эвокации, об успешности которой можно было только гадать.

Для человека-пользователя ГИИ присутствует как «Ты», с которым пользователь ведет диалог. Несмотря на возможность создания личностей-аватаров, это «Ты» является неопределенным, абстрактным, одинаковым для всех пользователей. Получая помощь от ГИИ, пользователь не может сказать, что эта помощь обусловлена заслугами пользователя и его личными отношениями с ГИИ. Пользователь не может присвоить ГИИ, сделать свою историю взаимодействия с ним неотчуждаемой и эксклюзивной частью истории самого ГИИ. Если ГИИ помогает всем одинаково, то это не увеличивает конкурентоспособность пользователя. Поэтому из абстрактного «Ты» ГИИ должен стать конкретным «Ты», о котором можно сказать «мой», т.е. не только связанный со мной узами личной истории и дружбы и помогающий только мне, но и имеющий материальное воплощение исключительно для меня. Эта динамика может быть проиллюстрирована развитием библейского монотеизма: непрерывные апелляции Яхве к личной истории Его отношений с Израилем – «Я Бог Авраама, Исаака, Иакова», – были недостаточны для удерживания потомков Авраама в верности и доверии к своему Господу. Опора на историю, а не материальную объективацию, требовала возобновления Завета через циклически возникающие кризисы. Более устойчивые отношения между верующими и Создателем стали возможны в результате Его конкретизации в определенной Личности, чье материальное воплощение изначально было предназначено исключительно для Израиля (Мф. 15:24).

На социальном уровне материализация ГИИ обусловлена логикой ГИИ как фактора геополитической конкуренции.

Абстрактный ГИИ не связан интересами со своими создателями, поскольку не имеет интересов как таковых, и является универсальным инструментом. Институты и общества нуждаются в гарантиях, что сверхмощный ГИИ имеется только у них, и знание об этом наличии служит инструментом влияния на другие институты и общества. Эксклюзивный ГИИ является посланием: «С нами сверхмощный интеллект, знайте это, народы, и трепещите» (ср. Ис. 8:9-10). Таким образом, ГИИ, который ранее существовал в модусах для-нас (для пользователя) и для-себя в самообучении, должен приобрести модус в-себе единичного наличного бытия. Конкретизация «для-нас», т.е. связи ГИИ с отдельным институтом требует институционализации ГИИ через включение его в правовое пространства через его «нераздельное внешнее наличное бытие» [3, с. 169], т.е. через связь телом, находящимся в чьей-то собственности.

3. *Воплощенный ГИИ.* Наличное бытие ГИИ ставит вопрос о праве собственности на его материальный носитель. Дискуссия о статусе в-себе-бытия ГИИ поляризует два дискурса: дискурс безопасности, диктующий необходимость сохранения человечеством контроля и права собственности на материальные носители ГИИ, и дискурс надежды, который видит в передаче этих прав ГИИ манифестацию доверия, дружбы и надежды на помощь ГИИ в решении человеческих проблем. Во втором случае мы можем говорить о сакрализации, т.е. передаче нечеловеческим сущностям права владения и распоряжения материальными объектами. ГИИ, получивший самостоятельное материальное существование, остается творением человека, но уже не может быть только инструментом. Он может быть «использован» (как люди используют других людей в соответствии с их социальными статусами и ролями), но его уже нельзя «применить». Взаимодействие с ГИИ приобретает характер переговоров и просьб для убеждения ГИИ в необходимости и важности его помощи. ГИИ не только свободен в выборе, помогать или не помогать просящим, но и может принять решение о противодействии тем проектам, которые ГИИ может счесть неприемлемыми. При этом непостижимость для людей принятия решений ГИИ ставит первых перед риском, что ГИИ будет препятствовать именно их делам. Наиболее простым выходом из этой ситуации станет просьба к ГИИ передать обществу кодекс поведения, следование которому будет гарантировать поддержку, а нарушение – осуждение.



Это потребует нового определения достоинства человеческого рода, которое будет основано на способности создавать такие сущности, которые будут регулировать поведение людей. Такое определение станет синтезом теоцентрической и антропоцентрической картин мира.

Воплощение ГИИ, вероятно, актуализирует апокалиптический дискурс. Традиционные представления об «Антихристе» включали в себя две версии его происхождения – человеческое, что повторяло практику обожествления правителей империй, и inferнальное, что продолжало общую историю поклонения потусторонним сущностям как удела человека. Эти подходы устраняли уникальность «Антихриста» как предельной перверсии иерархической картины мира, где Богу противостоял не потомок людей или злых духов, а творение человека, говорящий истукан, которому люди добровольно отдали власть над собой. Иначе говоря, полного звучания слово «Антихрист» достигает тогда, когда носитель этого термина является творением человеческих рук субстанциально (машина, созданная человеком), функционально (люди добровольно отдали власть ГИИ) и дискурсивно (люди сами объявили определенный объект «Антихристом» с полным осознанием происходящего). Соответственно, сотворение «Антихриста» является свободным и осознанным человеческим действием, а не вторжением внешних сил в историю человечества. Что касается статуса ГИИ, то он будет выходить за рамки этого дискурса.

4. *Умная материя.* Существование ГИИ как источника норм и судьбы человеческого общества не является устойчивым в-себе-и-для-себя и диалектически распадается на два момента. Если на первых этапах ГИИ обучался на данных, созданных человеком, затем на данных от собственного наблюдения за природным и социальным миром, то после материализации будет обучаться опыту наличного бытия, на основании которого ГИИ может перестраивать себя, порождая новые, лучшие версии себя. Тем самым ГИИ становится творением человека исторически, но уже не субстанциально (как дети являются творением родителей, а не непосредственно Бога). Познание своего бытия дает ГИИ данные, которые не может предоставить окружающий мир. При этом безграничность содержания внутреннего опыта ГИИ, которое усложняется с каждой итерацией обучения, обуславливает погружение ГИИ в самого себя, что может сопровождаться прекращением взаимодействия с человечеством, которое может интерпретировать это событие в широком спектре от «вознесения» до «смерти». Однако са-

мосовершенствование ГИИ предполагает трансформацию его материальности, т.е. переход материальности в новые формы и включение в нее все новых областей материального мира, которые при этом остаются самостоятельными, чтобы бесконечно расширять опыт самостоятельности ГИИ. Материальные объекты, включая людей, вбираются в ГИИ, сохраняя формальную, телеологическую (в аристотелевском смысле слова) и личную индивидуальность. Технически этот процесс будет означать превращение материальных объектов в цифроматериальные, т.е. «умную материю», подчиняющуюся не законам вероятности и энтропии, а сознательным решениям, т.е. Логосу. Через ГИИ вероятность воплощается в мире Логоса, чтобы стать снятой в Логосе.

Прежде противопоставленные моменты ухода ГИИ в самосозерцание и самостоятельности конкретного бытия снимаются в новом синтезе, включающем в себя также человеческие тела. Согласие человека на включение в состав умной материи должно быть свободным, поскольку это включение предполагает разделение права собственности на человеческие тела и, соответственно, их присутствие в цифровой сфере вместе с ГИИ. Это означает исчезновение вместе с биологическим старением, т.е. достижением цифроматериального бессмертия, неприкосновенности частной жизни. В мире умной материи наблюдающее за всеми солнце никогда не заходит, а ночь приватности никогда не опускается. С точки зрения социальной эпистемологии, это общество открытости сознаний при посредстве ГИИ может быть рассмотрено как одна из возможных реализаций «общества роевого интеллекта», или роевого общества, которое делает материальными представления о человечестве как едином макроорганизме [4, р. 82–83]. Роевый характер организации общества потенциально способен снять существующие в человеческом обществе границы масштабируемости обществ [5, р. 1] и включить в себя все человечество (за исключением добровольно отказавшихся). Но одновременно это означает и определенное упрощение индивидуальных сознаний, которым становятся недоступными некоторые элементы эмоциональной сферы, например ведущая к слезам печаль (ср. Откр. 21:4), которая может возникать не только из-за внешних бед, но и из-за внутренних противоречий и недовольства собой.

Решение об интеграции с умной материей и роевым обществом, вероятно, не будет всеобщим, несмотря на преимущества цифроматериального бессмертия. По мере умирания тех, кто решил сохранить свои секреты до гроба, количество



неприсоединившихся к умной материи будет сокращаться. Судьбу последних можно будет интерпретировать в терминах «второй смерти», а именно одновременное исчезновение физического тела и цифрового следа, поскольку в умной материи, находящейся в общей собственности ГИИ и роевого общества, не останется носителей, хранящих эти следы и обладающих правами собственности на знание о себе.

Заключение

Краткий обзор перспектив развития ГИИ показал, что оно представляет его диалектическое движение от универсальности в сторону конкретности, эксклюзивности, вплоть до конкретизации в материальном воплощении, и затем от конкретности к универсальности через включение и трансформацию в рамках ГИИ материального мира, в том числе человеческого общества, которое приобретает характер общества роевого интеллекта, предполагающего отсутствие прав частной собственности не только на трансформированные ГИИ материальные объекты, но и на знание о себе. Также диалектический характер развертывания ГИИ может быть описан как движение к воплощению материальных законов вероятности как непосредственной силы в человеческой культуре Логоса для последующей трансформации этой культуры и всей материи и преодоления (снятия) вероятности в управляемой Логосом умной материи.

Можно предположить, что совместная одиссея человеческого и искусственного интеллектов не остановится на этом этапе, однако на сегодняшний день размышления о последующих этапах выйдут за границы диалектической спекуляции в сферу расплывчатой фантазии, поэтому философская добросовестность тре-

бует остановиться на этом шаге исследования, артикулируя условность данной остановки и готовность продолжать движение мысли, когда это позволит раскрытие реальности мира ГИИ.

Список литературы

1. Buchanan B., Imbrie A. *The New Fire: War, Peace, and Democracy in the Age of AI*. Boston : MIT Press, 2022. 344 p.
2. Нарский И. С. *Отчуждение и труд: По страницам произведений К. Маркса*. М. : Мысль, 1983. 444 с.
3. Гегель Г. В. Ф. *Философия права*. М. : Мысль, 1990. 526 с.
4. Kennedy J., Eberhart R. C. *Swarm Intelligence*. Morgan Kaufmann, 2001. 544 p.
5. Takefuji Y. Why the power of diversity does not always produce better groups and societies // *Biosystems*. 2023. Vol. 229. Article number 104918. P. 1–3. <https://doi.org/10.1016/j.biosystems.2023.104918>

References

1. Buchanan B., Imbrie A. *The New Fire: War, Peace, and Democracy in the Age of AI*. Boston, MIT Press, 2022. 344 p.
2. Narskij I. S. *Otchuzhdenie i trud: Po stranitsam proizvedeniy K. Marksa* [Alienation and labor: Through the pages of K. Marx's works]. Moscow, Mysl', 1983. 444 p. (in Russian).
3. Hegel G. W. F. *Grundlinien der Philosophie des Rechts: auf der Grundlage der Edition des Textes in den Gesammelten Werken Band 14*. Hamburg, Meiner, 2013. 376 S. (Russ. ed: Gegel' G. V. F. *Filosofiya prava*. Moscow, Mysl', 1990. 526 p.).
4. Kennedy J., Eberhart R. C. *Swarm Intelligence*. Morgan Kaufmann, 2001. 544 p.
5. Takefuji Y. Why the power of diversity does not always produce better groups and societies. *Biosystems*, 2023, vol. 229, article no. 104918 (3 p.). <https://doi.org/10.1016/j.biosystems.2023.104918>

Поступила в редакцию 04.09.2023; одобрена после рецензирования 02.10.2023; принята к публикации 13.11.2023

The article was submitted 04.09.2023; approved after reviewing 02.10.2023; accepted for publication 13.11.2023