ФИЛОСОФСКИЕ ПРОБЛЕМЫ НАУЧНО-ТЕХНИЧЕСКОГО РАЗВИТИЯ

УДК 001 DOI: 10.31249/scis/2024.02.03

Зубковская А.А.*

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И НАУЧНАЯ ОБЪЕКТИВНОСТЬ

Zubkovskaya A.A.

ARTIFICIAL INTELLIGENCE AND SCIENTIFIC OBJECTIVITY

Аннотация. Статья подготовлена в рамках дискуссии на круглом столе по проблеме использования искусственного интеллекта в научной деятельности 18 января 2024 года в Институте научной информации по общественным наукам (ИНИОН) РАН. В настоящей работе осуществляется постановка проблемы научной объективности в контексте эпистемологии добродетелей. Объективность рассматривается в связи с понятием субъекта и некоторыми его философскими интерпретациями. В рамках эпистемологии добродетелей научная объективность истолковывается как практикуемая субъектом познания добродетель, исходящая от субъекта и налагающая на него моральную ответственность за научную деятельность. В такой перспективе объективность обладает чертами праксиса, поэтому не может практиковаться искусственным ин-

* *Зубковская Анастасия Александровна* – кандидат философских наук, научный сотрудник Института научной информации по общественным наукам (ИНИОН) РАН, Москва, Россия; zubkovskaya@inion.ru ORCID 0000–0002–6232–6837

Zubkovskaya A.Al. – PhD, scientific researcher, Institute of Scientific Information for Social Sciences of the Russian Academy of Sciences, Moscow, Russia; zubkovskaya@inion.ru ORCID 0000–0002–6232–6837

теллектом, который не принимает на себя моральную ответственность за результаты своей деятельности. В связи с этим природа ИИ интерпретируется не как «расширенное сознание», а как инструмент, появившийся в результате эволюции человеческой культуры. В заключении приведены рассуждения о том, что для трактовки понятия объективности необходим субъект, обладающий телесностью и принимающий на себя этические обязательства за свои действия, поэтому требования научной объективности для искусственного интеллекта являются бессмысленными, а сам ИИ не может стать полноправным актором научно-исследовательского процесса, однако является полезным инструментом для улучшения его эффективности.

Ключевые слова: искусственный интеллект; субъект; объективность; добродетели; эпистемология; эволюционизм.

Abstract. This paper has been prepared as a part of a round table discussion on the problem of using artificial intelligence in science on January 18, 2024 at the Institute of Scientific Information for Social Sciences of the Russian Academy of Sciences. In this article, the problem of scientific objectivity is posed in the context of the virtue epistemology. Objectivity is considered in connection with the concept of the subject and some of its philosophical interpretations. Within the framework of the epistemology of virtues, scientific objectivity is interpreted as a virtue practiced by the subject of knowledge, emanating from the subject and imposing moral responsibility for scientific activity on the subject. In this perspective, objectivity has the features of praxis, so it cannot be practiced by artificial intelligence, which does not assume moral responsibility for the results of its activities. In this regard, the nature of AI is interpreted not as an “extended mind”, but as a tool that appeared as a result of the evolution of human culture. In conclusion, the reasoning is given that in order to interpret the concept of objectivity, a subject with body and assuming ethical obligations for his actions is necessary, therefore, the requirements of scientific objectivity for artificial intelligence are meaningless, and AI itself cannot become a full-fledged actor in the research process, however it is a useful tool for improving its effectiveness.

Keywords: artificial intelligence; subject; objectivity; virtues; epistemology; evolutionism.

Введение

Чем больше современный человек взаимодействует с искусственным интеллектом, тем больше возникает теоретических и практических проблем, связанных с внедрением последнего в различные сферы жизни общества. Вокруг использования ИИ ведутся оживленные дискуссии в публичной сфере и в рамках профессиональных сообществ, в том числе среди представителей академической среды (см., например: [Бажанов]).

После череды ставших широко известными случаев использования ChatGPT при написании квалификационных работ и научных статей в академическом сообществе обострилась проблема этического и правового регулирования обращения к искусственному интеллекту в научной деятельности. В контексте этой дискуссии вполне правомерен философский вопрос о сравнении результатов интеллектуального труда ученого и нейросети. То есть, справедливо ли утверждать, что искусственный интеллект является более объективным, нежели человек с его «idola»?

В этой работе предлагается рассмотреть сформулированную проблему в свете высказывания «объективность – это эпистемическая добродетель», отсылающего к дискурсу эпистемологии добродетелей – философскому подходу, соединяющего эпистемологические и ценностно-нормативные этические категории.

Эпистемология добродетелей

Как отмечает И.Т. Касавин, одной из характеристик эпистемологии добродетелей (virtue epistemology, VE) является субъектоцентризм, т.е. «убеждение в том, что индивидуальные и коллективные субъекты познания, люди и сообщества составляют главный источник эпистемических ценностей и предмет эпистемической оценки» [Касавин, с. 8]. Он указывает, что под добродетелями в этом направлении могут пониматься как когнитивные способности, так и моральные характеристики, если они имеют приставку «эпистемический».

Определение природы и характера эпистемических добродетелей – дискуссионная проблема, однако если смотреть на этот феномен с позиции «респонсибилизма» (virtue responsibilism, англ.

responsibility – ответственность), то, согласно этому подходу, эпистемические добродетели не только характеризуют уровень развития когнитивных способностей субъекта (интуиция, дедукция и т.д.), но также обладают нормативной силой, наделяют субъекта познания моральной *ответственностью* за результат своей деятельности [Касавин, с. 11].

«Субъект является источником интеллектуальных добродетелей» [Касавин, с. 12]. Для эпистемологического анализа это означает, что интеллектуальные добродетели не являются врожденными характеристиками, но приобретаются в процессе научения. Так, если бы объективность была просто когнитивной чертой, мы бы могли приписать ее любым сознаниям в сравнительной перспективе, однако она является частью этоса науки, который, при этом, не является набором абстрактных норм, а формируется в интеллектуальной истории [см.: Аргамакова, Махаев, с. 51].

Субъект

В исследовательской литературе отмечается, что «конструктивистская эпистемология наиболее выразительно показывает условия и следствия подмены объективности интерсубъективностью, а предметного знания – социальной конструкцией <…> В этом смысле границы между субъективным и объективным, социальным и физическим размывались до исчезновения» [Аргамакова, Махаев, с. 48]. Так что, говоря в антилатуровском ключе, что субъектом интеллектуальной истории является человек, мы имеем в виду, что нечеловеческие участники познавательного процесса выполняют роль инструментов, а не самостоятельных акторов этой деятельности.

Однако и в этом ключе категорию «субъект» также можно подвергнуть критике, как, например, это происходит в теории активного экстернализма (active externalism) Кларка и Чалмерса. «Где заканчивается сознание и начинается весь остальной мир?» – спрашивают авторы тезиса расширенного сознания (extended mind thesis) [Clark, Chalmers, p. 7]. Согласно этому тезису, интеллектуальные артефакты (например, записные книжки или мобильные гаджеты) являются «внешней когнитивной подсистемой субъекта» [Сущин, с. 173].

М.А. Сущин, критикуя тезис расширенного сознания в своем докладе на упомянутом круглом столе и в своих публикациях, указывает, что «субъект – это не некий абстрактный вычислительный код, который можно реализовать неограниченное число раз на разном “железе” <…> Субъект обладает определенной телесностью» [Сущин, с. 171].

Экстерналистский подход в данном случае сам по себе любопытен и приводит к некоторым теоретическим наблюдениям в области эволюции человеческого разума и использования искусственного интеллекта. У Кларка и Чалмерса читаем: «Определенно кажется, что эволюция благоприятствовала встроенным возможностям, которые особенно приспособлены для паразитирования на локальной среде, чтобы уменьшить нагрузку на память и даже изменить природу самих вычислительных задач» [Clark, Chalmers, p. 11]. Можно согласиться с тем, что культурно-эволюционная траектория человека может рассматриваться как «экстернальная», иными словами, эволюционные адаптации человека чаще всего являются изобретением или функциональным приспособлением, взятым из внешней среды. Так, например, если шерсть млекопитающих как эволюционная адаптация для поддержания терморегуляции все же является частью их тела, то одежда человека, выполняющая в том числе такую же функцию сохранения тепла, не является продолжением его телесности. Другой иллюстрацией могут быть зубы и когти хищников, в то время как такую же функцию у человека выполняют орудия (например копья или ножи), не являющиеся его продолжением. Культурная эволюция, таким образом, может быть охарактеризована в «пассивном» экстерналистском ключе. Этот подход можно назвать экзоэволюционным.

Как интерпретировать с этой точки зрения появление искусственного интеллекта? Как инструмент или адаптацию. Однако, так же как мы не считаем микроскоп или очки частью своего собственного тела, нейросеть не является «продолжением» нашего сознания или самостоятельно действующим актором, потому что она функционирует в момент нашего к ней обращения. Иными словами, искусственный интеллект не выполняет роль субъекта самостоятельного или расширенного, и, таким образом, если мы применяем к ИИ понятие научной объективности, то такое же требование можно применить к любым научным приборам.

Самость

Л. Дастон и П. Галисон исследовали научную объективность как эпистемическую добродетель в контексте понятия самости (self) [Дастон, Галисон]. На мой взгляд, это более удачный термин для идентификации ученого – того, кто управляет своей научной самостью, но не перестает быть субъектом познания. С точки зрения авторов, «устройство самости исторически менялось, и разные ее версии – к примеру, конгломерат способностей, динамическая субъективность, выстроенная вокруг воли, или самость, зажатая между реальностью и бессознательным, – сильно отличались друг от друга» [Дастон, Галисон, с. 21].

Они объясняют, что появление объективности во многом связано с интерпретацией пары «субъективный и объективный» в философии Иммануила Канта, который внес большой вклад в то, чтобы смысл терминов поменялся с точностью до наоборот, поскольку в классической философии до этого субъект и объект означали полную противоположность тому, как они истолковываются нами сейчас [Дастон, Галисон, с. 306]. Со временем «объективная значимость» выходит на первый план в истории науки, формируя новый тип научной самости – объективного ученого, чья беспристрастная деятельность исключает вмешательство субъективных свойств в исследовательский процесс. Более того, в какой-то момент научная объективность становится неотъемлемой частью классического понятия самой науки.

А.В. Баева отмечает, что «создание объективного образа требовало соблюдения ограничений в отношении научной самости» [Баева, 2021, с. 108]. Размышляя над книгой Дастон и Галисона, она использует термин «цифровая объективность» для обозначения возможных путей трансформации понятия научной объективности и допускает, что со временем, благодаря трансформациям в приемах визуализации, эпистемологическое напряжение между субъектом и объектом может быть переведено в такой режим, при котором мы сможем говорить о появлении цифровой объективности как новой эпистемической добродетели [Баева, 2021, с. 110].

Заключение

Несмотря на то, что дискуссии по вопросам использования ИИ в научной деятельности начались относительно недавно, уже можно констатировать, что искусственный интеллект *расширяет* исследовательские навыки ученого, но *не заменяет* их. Нельзя просто сказать, что ИИ более или менее объективен, поскольку он – не субъект и не обладает субъективными качествами и состояниями. Понятие субъекта и объекта *взаимно* определяют друг друга.

Что значит быть субъектом? Этот вопрос имеет теоретическую глубину, которая требует продолжения философских исследований и дискуссий. Тем не менее, если присвоить понятию субъекта такие неотъемлемые характеристики, как например телесность, то можно говорить о том, что существуют разные типы субъектов, поскольку их телесность тоже разная. Можно ли в таком случае предъявлять человеческое требование объективности, которое рождается в рамках интеллектуальной истории, например к животным? В этом плане объективность приобретает черты праксиса для того, чтобы осуществлять и осмыслять научную деятельность, поскольку та интерпретируется тоже в практическом ключе: не «что есть наука?», а «как она делается?» [Баева, 2023, с. 37]. Я бы добавила сюда вопрос «кем она делается?».

Кроме того, научная объективность, интерпретируемая с точки зрения эпистемологии добродетелей как эпистемическая практика, исходит от субъекта, управляющего своей самостью, и налагает на него моральную ответственность за результат его научного труда. Такой взгляд позволяет ступить на территорию постмодернистской критики авторства и поставить вопрос о теоретической реабилитации и удержании автора как субъекта научной деятельности, но уже с учетом новых эпистемологических траекторий.

Список литературы

Аргамакова А.А., Махаев М.Р. Проблема объективности в соотношении с методом социальной эпистемологии // Вестник КНИ РАН. Серия Социальные и гуманитарные науки. – 2022. – № 3 (3). – С. 45–63.

Баева А.В. Современные исследования «науки в действии»: знание с точки зрения эмпирической эпистемологии // Вестник Московского университета. Серия 7: Философия. – 2023. – № 1. – С. 35–49.

Баева А.В. Трансформация научной объективности и научной самости: особенности этико-эпистемологического режима цифровизации // Вопросы философии. – 2021. – № 1. – С. 104–114.

Бажанов В.А. Искусственный интеллект и технологии BIG Data (большие данные). Тернистый путь сквозь туман метафор // Философия. Журнал Высшей школы экономики. – 2023. – Т. 7, № 3. – С. 251–256.

Дастон Л., Галисон П. Объективность. – Москва: Новое литературное обозрение, 2018. – 577 с.

Касавин И.Т. Эпистемология добродетелей: к сорокалетию поворота в аналитической / Эпистемология и философия науки. – 2019. – Т. 56, № 3. – С. 6–19.

Сущин М.А. Субъект: воплощенный, встроенный в среду, но не расширенный // Субъективный мир в контексте вызовов современных когнитивных наук / под ред. В.А. Лекторского, Е.О. Труфановой, А.Ф. Яковлевой. – Москва: Аквилон, 2017. – С. 170–181.

Clark A., Chalmers D. The extended mind // Analysis. – 1998. – Vol. 58, N 1. – P. 7–19.

References

Argamakova A.A., Mahaev M.R. The problem of objectivity in relation to the method of social epistemology // Vestnik KNI RAN. «Social science and Humanities». 2022. – N 3 (3). – Pp. 45–63. (In Russian)

Baeva A.V. Modern research of the “science in action”: knowledge from empirical epistemology point of view // Bulletin of the Moscow University. Philosophy series, 2023. – N 1. – Pp. 35–49. (In Russian)

Baeva A.V. Transformation of scientific objectivity and scientific self: specificities of ethical and epistemological regime of digitalization // Problem of Philosophy. – 2021. – N 1. – Pp. 104–114. (In Russian)

Bazhanov V.A. Artificial intelligence and Big data technology. The thorny path through the fog of metaphors // Philosophy. Journal of the Higher School of Economics. – 2023. – Vol. 7, No 3. – P. 251–256. (In Russian)

Clark A., Chalmers D. The extended mind // Analysis. – 1998. – Vol. 58, N 1. – P. 7–19.

Daston L., Galison P. Objectivity. M.: New Literary Review, 2018. – 577 p. (In Russian)

Kasavin I.T. Virtue epistemology: on the fortieth anniversary of analytical philosophy // Epistemology and philosophy of science. – 2019. – Vol. 56, N 3. – Pp. 6–19. (In Russian)

Sushchin M.A. Subject: embodied, embedded in the environment, but not extended] // The subjective world in the context of the challenges of modern cognitive sciences/ V.A. Lektorskiy, E.O. Trufanova, A.F. Yakovleva (eds.). – M.: Akvilon, 2017. – P. 170–181. (In Russian)