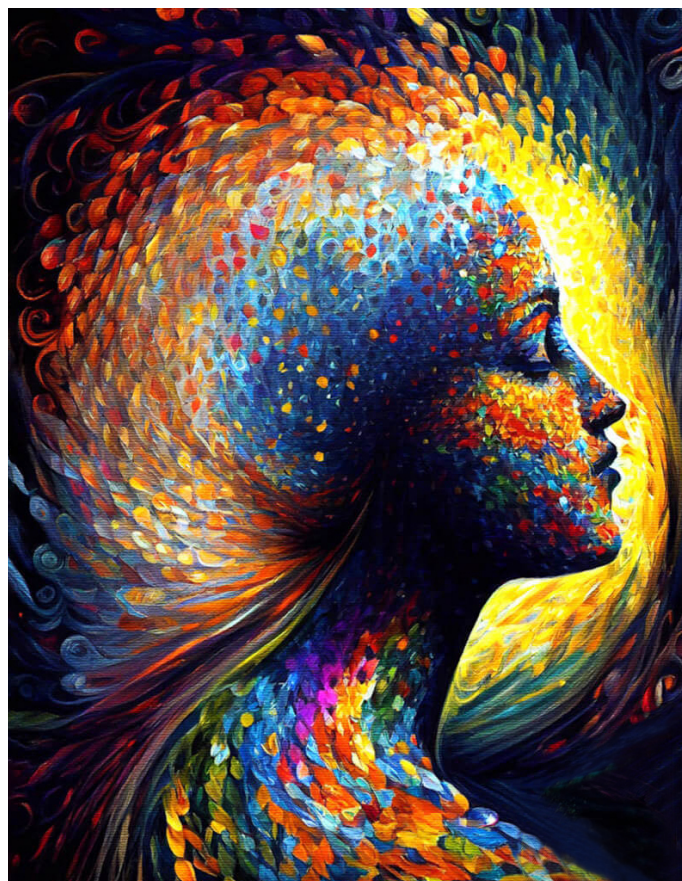


Искусственный интеллект — не интеллект и никогда им не будет



Игорь Прись,
ведущий научный
сотрудник Института
философии НАН Беларуси,
кандидат физико-
математических наук,
доктор философии (PhD)

Так называемый искусственный интеллект (ИИ) не мыслит. Во всяком случае — так, как это делают люди, если не ограничиваться чисто механическими формами, например употреблением таблицы умножения или логических правил. Он не способен к рефлексии, то есть к мысли о мысли. У него нет сознания (причем не только феноменального) и интенциональности, в частности отсутствуют убеждения и желания. ИИ не рационален и не иррационален, не способен самостоятельно формировать планы, ставить перед собою цели и задачи.



Естественный интеллект (ЕИ) отличается материальностью (человек — биологическое существо, и это принципиально), а также тем, что он способен пользоваться языком, который не может быть редуцирован к математическим алгоритмам, механическому следованию нормам. ИИ следует формальным правилам, а мы — естественным, укорененным в индивидуальной или социальной практике, опыте, реальности. Их применение чувствительно к контексту, не предопределено, но в то же время рационально. ИИ лишен гибкости и универсальности. У человека, в отличие от машины, есть свобода воли.

Существующий ныне ИИ называют «слабым», это лишь имитация интеллекта человека. Но говорят также о гипотетическом, «сильном» или «прометеевском» ИИ, который действительно был бы способен мыслить по образу и подобию ЕИ и, быть может, ощущать и чувствовать. Существует несколько широко известных аргументов против сильного ИИ: аргумент «китайская комната» американского философа сознания и ИИ Дж. Сёрла, аргумент с привлечением теоремы Гёделя о неполноте, аргумент американского философа Х. Дрейфуса о противоречивости самого

проекта символического сильного ИИ и др. [1]. На мой взгляд, все они эксплицитно или имплицитно основываются на принятии категориального различия между идеальным и реальным. Человек и его сознание реальны. Написанная им для машины программа идеальна. Не имеет смысла трактовать идеальное как реальное. В другом важном смысле можно сказать обратное: работа ИИ (компьютера) отличается от мыслительной деятельности человека тем, что первая материальна, тогда как мысль идеальна. Нельзя трактовать реальное как идеальное. С другой стороны, искусственный интеллект не функционирует сам по себе, а употребляется людьми. Достаточно высокая степень интеграции человека и машины позволяет говорить о расширенном интеллекте первого, который как бы одухотворяет второе.

ИИ, трансгуманизм и нейроцентризм

Современный материализм утверждает, что человек и его сознание редуцируются к его мозгу, нейрональным процессам. Трансгуманизм постулирует, что человек – это программа, которая может быть реализована и на другом материальном носителе. Компьютерный бог был бы просто очень мощной программой, написанной после момента так называемой сингулярности, идею которой в 1965 г. выдвинул британский математик И. Дж. Гуд: «Определим сверхумную машину как такую машину, которая может значительно превзойти все интеллектуальные активности сколь угодно умного человека. Поскольку проектирование машин – одна из этих интеллектуальных активностей, сверхинтеллект сможет спроектировать еще лучшие машины; в результате неизбежно произойдет «интеллектуальный взрыв», и интеллект человека окажется далеко позади» [2].

Трансгуманизм основан на предположении, лежащем в основе рассуждения Гуда, что ИИ способен превзойти интеллект человека. Например, австралийский философ Д. Чалмерс допускает возможность сингулярности и рассматривает 4 варианта места человека в трансгуманистическом мире: уничтожение, изоляция, неполноценность, интеграция [3]. Поэтому он призывает встраивать в искусственный интеллект подходящие ценности и поначалу создавать машины ИИ в виртуальном, а не реальном мире.

На самом деле это антропоморфный взгляд, проецирующий характеристики личности и общества на совсем другой вид реальности. О невозможности

создания машины умнее человека говорил советский философ Э.В. Ильенков [4]. Многие современные исследователи считают, что когнитивные процессы индивидуума – не алгоритмы и не могут быть автоматизированы. Принципиальную возможность существования сверхумной машины также отрицает неокзистенциализм немецкого философа М. Габриэля на том основании, что, в отличие от человека, у компьютера нет и не может быть духа, то есть он не способен рефлексировать. ИИ – это, так сказать, интеллект первого порядка – механически выполняющий логические операции [5]. У нас есть радикальная свобода в выборе ответа на вопрос о том, кто мы есть. То есть человек – это рациональное «существо, которое ведет свою жизнь в свете концепции того, что значит быть человеческим существом» [6]. Поэтому никакой суперкомпьютер не может превзойти человека. Нейроцентризм, включающий в себя нейрофизикализм, Габриэль называет нейроманией, а французский философ Ж. Бенуа – нейроидеализмом. Габриэль также видит политическую подоплеку идеологии натурализма: «Натурализм – <...> упрощенная материалистическая идеология, изначально продвигавшаяся американскими идеологическими государственными аппаратами (включая университеты) в целях борьбы с советскими идеями» [5]. Французский философ науки М. Битболь считает, что метафизический натурализм лежит в основе западной философии модерна и что сейчас наступает эпоха не постгуманизма, а постнатурализма [7].

На наш взгляд, интеллект – нормативный концепт, а дух, о котором говорит Габриэль, – нормативность, компьютер же выполняет операции механически. Также разумно допустить, что интеллект предполагает наличие сознания и феноменального сознания. Но, как и М. Битболь, мы полагаем, что «искусственное (феноменальное) сознание – бессмысленный концепт» [7]. Это так, поскольку между нашим непосредственным перцептивным (чувственным) опытом и реальностью нет никакой дистанции (это не перцептивная вуаль модерна). Для известного британского философа Т. Уильямсона «суть сознания в том, чтобы включить мир» [8]. В частности, знание – ментальное состояние в экстерналистском смысле. Поэтому сознание и интеллект – не чисто нейробиологические процессы внутри черепной коробки, и по этой же причине ИИ ничего не может знать о реальности. Искусственный и человеческий – два разных вида интеллекта. Как пишет французский философ когнитивных наук Д. Андлер, «нельзя сравнивать ковбоя и его тень. Тень зависит от ковбоя и лишь смутно напоминает его» [9].

Интеллект как нормативность

Габриэль употребляет термин «дух» вместо зонтичного понятия «сознание» в смысле английского mind. Это призвано отразить тот факт, что различие между сознанием и природой категориальное. Сознание как дух – не природный вид, а «объяснительная» структура, позволяющая отличить человеческое существо одновременно от неодушевленной природы и других представителей животного мира [6]. Для Габриэля дух связан с действием. Это позволяет избежать овеществления первого.

На самом деле такая точка зрения оказывается ограниченной, поскольку очевидно, что дух связан не только с действием, но и с мыслью и смыслом, которые могут существовать независимо от него. Но дух, как было отмечено, понимается нами как нормативность [10, 11]. Он идеален. Это значит, что термин «дух» не является жестким десигнатором (не отсылает во всех возможных мирах к одной и той же сущности), то есть сознание как дух вообще не вид, а не только не естественный вид; оно не имеет сущности. Принять во внимание нормативность – значит осознать категориальное различие между идеальным и реальным и, соответственно, различие между видимостью и реальностью. То, что обладает духом, не редуцируется к тому, что оно просто есть – своей онтологии. Натурализм преодолевается. Бенуа пишет: «Вещи есть то, что они есть, – это их определение. С другой стороны, мы можем ошибаться в том, что они есть, и даже можем ошибаться в том, что мы есть (...)». Существование в соответствии с нормой (...) не то же самое, что просто «существование» [10].

Таким образом, по мнению автора, гипотеза, что ИИ может превзойти интеллект человека или хотя бы приблизиться к нему, как она ставится в рамках современной редуктивной натуралистической парадигмы, идеалистична. ИИ – не интеллект и никогда им не будет. ЕИ отличается чувствительностью к контексту, наличием духа, который идеален. В то же время, что касается решения конкретных алгоритмизируемых проблем, то пока не видно пределов возможного прогресса ИИ. Этот вывод подтверждается точкой зрения Д. Андлера: «Интеллект – это не вещь, не явление, не процесс и не функция, а норма, которая применяется к поведению: он квалифицирует отношение между человеком и его миром, причем таким образом, который никогда не является объективным и окончательным...» [9]. «Искусственный интеллект, в какой бы форме и на каком бы уровне развития он ни находился, предназначен для решения проблем, что явля-

ется лишь второстепенной задачей для человеческого интеллекта» [9].

Известно, что поведение людей зачастую подчиняется квантовой, а не классической логике [12]. Это объясняется тем, что последняя в известном смысле принимает во внимание свойственные решениям и поступкам людей не(пред)определенность и контекстуальность. Можно поэтому предполагать, что ИИ, основанный на квантовой логике (квантовой теории информации, но необязательно на квантовой физике), будет человекоподобным. Но и он, согласно нашему аргументу, никогда не станет умным, не приблизится к ЕИ.

Сказанное, как кажется, не исключает принципиальную возможность интегрирования в нашу форму жизни нечеловеческих, быть может, искусственных или инопланетных существ, которые могли бы рассматриваться как обладающие интеллектом и сознанием. В настоящее время, однако, не только не существует критериев, которые позволили бы установить наличие у таких гипотетических существ сознания и интеллекта, но даже неясно, каким образом такие критерии могли бы быть выработаны. То есть их сознание и интеллект оказываются недоступными и метанедоступными.

ИИ, большие данные и инфосфера

Классический символический ИИ – это программа, алгоритм, расширенная логика. Сменивший его коннекционистский ИИ – искусственная нейронная сеть. Философия первого – рационализм («всё – логика!»); философия второго – эмпиризм («всё – перцепция!»), хотя он включает в себя и существенные элементы символического ИИ. Так называемое глубокое обучение и большие языковые модели (chatGPT и др.) – современное развитие коннекционизма. Предположительно, ИИ ближайшего будущего объединяет оба подхода. Парадигму такой гибридизации можно условно сравнить с критической философией Канта, синтезирующей рационализм и эмпиризм.

Коннекционистский ИИ достиг впечатляющих успехов. Он (но не в одном лице) общается с человеком, проходит тест Тьюринга, пишет научные статьи, переводит с одного языка на другой, управляет автомобилем, распознает объекты и образы и т.д. Это стало возможным благодаря научно-техническим достижениям и появившейся способности накапливать и обрабатывать большие данные.

Экспоненциальный рост инфосферы, граница между которой и ИИ размывается, способствует возникновению многочисленных практических,

этических и эпистемологических проблем. К разряду последних относятся, например, вопросы экспертизы и научного метода, референтного класса, редкого (или нового) паттерна, онтологической окклюзии и «опасного знания» (профилирование, микротаргетирование и др.) [13]. К этическим – проблемы, связанные со свободой мысли в ситуации расширенного познания, правом собственности на данные, их приватностью или же, наоборот, транспарентностью, монетизацией, согласием на их использование, информационной зависимостью и т.д. [13, 14].

На вопрос о том, теряем ли мы контроль над цифровыми технологиями и инфосферой, Э. Ли, специалист в этой области, отвечает так: «У нас никогда не было этого контроля. Нельзя потерять то, чем не обладают» [15]. Можно, таким образом, говорить о возникновении новой реальности – цифровой. Это понятие обобщает онтологию документальности М. Феррариса [16]. Большие данные, и в частности Интернет – не представление мира, а сами по себе особый вид реальности, пространство регистрации [17–20]. Цифровая действительность предполагает природную, социальную и другие виды реальности и их законы. В то же время, будучи новым видом, она требует своих специфических экспертизы и научных методов.

Эмпирический подход к ИИ породил версию радикального эмпиризма, согласно которой любое знание может быть получено в результате более или менее сложной формы индукции, а наука будущего не станет нуждаться в теории. На самом деле использование больших данных не отменяет необходимость в гипотезах, моделях, теориях, объяснениях [19, 20]. Без развития научного метода применительно к первым их накопление оказывается просто расширением области непознанной реальности, увеличивающим относительное незнание. «Без математических или концептуальных моделей данные просто шум» [21]. В новой науке, центрированной на данных [22], большую роль играют теория вероятностей и статистический анализ. Многое зависит от умения видеть важное, ставить правильные вопросы и давать на них ответы, принимать верные решения [23]. Тезисы «корреляции достаточны», «при достаточном количестве данных числа говорят сами за себя» ошибочны.

Многие исследователи преследуют цель создания сильного ИИ, то есть автономного, по образу и подобию естественного. Эта задача, однако, идеалистична, недостижима (хотя на этом пути, как это бывает, возможны другие открытия). Никакой, сколь угодно мощный ИИ, даже если он будет способен решать практически любые (алгоритмизируемые)

проблемы гораздо быстрее и лучше человека, не сможет приблизиться к ЕИ. Как пишет Андлер о гипотетическом сильном ИИ, «искусственный интеллект такого рода (...) должен был бы отвечать противоречивым требованиям: быть человеческим – знать человеческую ситуацию как человек – и быть искусственным» [9]. В то же время прометеевский проект ИИ вызывает опасения, так как способен привести к возникновению псевдоавтономного ИИ – непонятного и неконтролируемого, лишаящего человека автономии. Он представлял бы не меньшую угрозу для человека, чем сверхинтеллект, воображаемый трагуманизмом. Не следует забывать, что ИИ – всего лишь инструмент, совершенствование которого следует направлять на благо человека. ■

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Bringsjord S. Artificial intelligence // <https://plato.stanford.edu/entries/artificial-intelligence/#StroVersWeakAI>.
2. I.J. Good. Speculations concerning the first ultraintelligent machine // *Advances in Computers*. 1965. Vol 6. P. 31.
3. D. Chalmers. The Singularity: A Philosophical Analysis // *Journal of Consciousness Studies*. 2010. №17. P. 7–65.
4. Ильенков Э.В. Об идолах и идеалах. / Э.В. Ильенков. – М., 1968.
5. Gabriel M. Neo-existentialism. How to conceive of the human mind after naturalism's failure / M. Gabriel. – Medford, 2018.
6. Gabriel M. *Propos réalistes* / M. Gabriel. – Paris, 2020.
7. M. Bitbol. La conscience artificielle: Une critique pensée et vécue // *Chroniques Phénoménologiques*. 2018. №10. P. 5–16.
8. T. Williamson. Acting on Knowledge // *Knowledge-First*. – Oxford, 2017. P. 163–181.
9. Andler D. Intelligence artificielle, intelligence humaine: la double énigme / D. Andler. – Paris, 2023.
10. Benoist J. Does mind «exist»? Comments on Markus Gabriel's «Neo-Existentialism: How to Conceive of the Human Mind After Naturalism's Failure» // <https://www.cairn.info/publications-deBenoist-%20Jocelyn-20377.htm>.
11. И.Е. Прись. Природа и сознание // *Вестник Челябинского университета*. 2018. №9. С. 18–23.
12. Khrennikov A. Open quantum systems in biology, cognitive and social sciences / A. Khrennikov. – Heidelberg, Berlin, New York, 2023.
13. Carter A. Digital knowledge. A philosophical investigation / A. Carter. – Routledge, 2023.
14. A. Carter A. Varieties of (extended) thought manipulation // *The Future of Freedom of Thought: Liberty, Technology, and Neuroscience*. 2024. №9.
15. E.A. Lee. Are we losing control? // *Perspectives on digital humanism*. Cham, 2022.
16. Ferraris M. Documentality. Why it is important to leave traces / M. Ferraris. – Fordham, 2012.
17. И.Е. Прись. Виртуальная реальность, искусственный интеллект и контекстуальный реализм // *Сибирский философский журнал*. 2021. Т. 19. №2. С. 158–180.
18. Benoist J. *New Realism in the World Picture Age* / J. Benoist. – Madrid, 2020.
19. И.Е. Прись. Эмпирические данные без теоретической нагрузки // *Философские исследования*. 2021. Вып. 8. С. 76–88.
20. Прись И.Е. Знание в контексте / И.Е. Прись. – СПб., 2022.
21. M. Pigliucci. The End of Theory in Science? // *EMBO Reports*. 2009. №10 (6). P. 534–34.
22. Leonelli S. Data-Centric Biology: A Philosophical Study / S. Leonelli. – Chicago, 2016.
23. L. Floridi. Big Data and Their Epistemological Challenge // *Philosophy and Technology*. 2012. Vol. 25. №4. P. 435–37.