

обработки информации. А значит, создание мыслящей машины вполне реально. Срединную позицию в этом вопросе занимает Х.Патнэм, высказывающийся против абсолютного уравнивания искусственного интеллекта и человеческого разума.

Современный этап развития концепции искусственного интеллекта опирается на процессы расширения нейросетей, развития цифрового обучения и генеративного искусственного интеллекта, что, бесспорно, обусловило возникновение множества футурологических концепций, таких как концепция трансгуманизма Н. Бострома и технологической сингулярности Р. Курцвейла. По мнению Н.Бострома, новые научные и цифровые достижения при разумном использовании имеют достаточные возможности для совершенствования физических и интеллектуальных возможностей человека. Особый оптимизм по отношению к цифровому будущему высказывает американский футуролог Р. Курцвейл, когда верит в технологическую сингулярность, когда в режиме синергии машины будут самосовершенствоваться, усиливая свой интеллект, превзойдя человеческий. Наступит постчеловеческая цивилизация с появлением

киборгов, цифрового (сознательного) бессмертия, радикального количественного и качественного улучшения человеческой жизни.

Таким образом, футуристических прогнозов на предмет роли искусственного интеллекта в жизни и судьбе человека и общества много, они противоречивы, от радужности до крайней тревожности. Вопрос в том, является ли искусственный интеллект мощнейшим человеческимобретением для преодоления цивилизационных вызовов или же мы обречены на враждебное сосуществование с ним, утрачивая человеческую природу. Ответ один – и то, и другое. Все зависит от целеполагания, присущее только человеку. Если цель благая – мы обретаем, в противном случае – мы обречены на ницшеанское «слишком человеческое»...

Список использованной литературы:

1. Зброжек Е.А. Онтология разрыва как ключ к пониманию философии С. Жижека // Вестник Воронежского государственного университета. Серия: Философия. – 2012. – № 2. – С. 115–120.

2. Декарт Р. Размышления о первой философии [Электронный ресурс]. – URL: [https://nibiryukov.mgimo.ru/nb\\_russian/nbr\\_teaching/nbr\\_teach\\_library/nbr\\_library\\_classics/nbr\\_classics\\_descartes\\_meditationes\\_de\\_prima\\_philosophia.htm](https://nibiryukov.mgimo.ru/nb_russian/nbr_teaching/nbr_teach_library/nbr_library_classics/nbr_classics_descartes_meditationes_de_prima_philosophia.htm) (дата обращения: 02.09.2025).

3. Гоббс Т. Левиафан [Электронный ресурс]. – URL: [https://www.civisbook.ru/files/File/Gobbs\\_Leviafan.pdf](https://www.civisbook.ru/files/File/Gobbs_Leviafan.pdf) (дата обращения: 02.09.2025).

4. Пушкарский А.Г. О судьбе центрального философского замысла создателя алгебры логики: к 200-летию Джорджа Буля // Вестник Балтийского федерального университета им. И. Канта. Серия: Гуманитарные и общественные науки. – 2015. – № 12. – URL: <https://cyberleninka.ru/article/n/o-sudbe-tsentralnogo-filosofskogo-zamysla-sozdatelya-algebry-logiki-k-200-letiyu-dzhordzha-bulya> (дата обращения: 03.09.2025).

5. Миронов Д.Г. Проблема интенциональности в философии Ф. Brentano // Философские науки. – 2013. – № 5. – С. 38–53.

УДК 130.2:004.8

## **ВОЗМОЖНОСТИ И РИСКИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: ФИЛОСОФСКИЙ РАКУРС ПРОБЛЕМЫ**

Калдыбаева Гулжан Зейнуллоевна

Карагандинский национальный исследовательский университет имени академика Е.А.Букетова,  
магистрант кафедры философии и теории культуры, Караганда, Казахстан  
(научный руководитель: Сейфуллина Галия Рустембековна  
кандидат философских наук, ассоциированный профессор)

Искусственный интеллект (ИИ) является одной из самых обсуждаемых тем на глобальной арене. Он активно внедряется в различные сферы жизни, от медицины и образования до промышленности и повседневной жизни. Этот процесс сопровождается как большими

возможностями, так и серьезными вызовами. В данной статье рассматриваются ключевые аспекты влияния ИИ на общество, а также риски и перспективы, которые он открывает.

Однако за рамки прикладного интереса к ИИ кроется глубинная философская проблематика, а именно вопрос о соотношении технической рациональности и человеческой свободы, о природе сознания, субъективности, морали и реальности. Мы попытаемся рассмотреть феномен ИИ в контексте философской традиции – от Платона и Канта до современных дебатов в аналитической философии и постгуманизме.

Искусственный интеллект — это система, способная выполнять задачи, традиционно ассоциировавшиеся с человеческой когнитивной активностью, такие как обработка естественного языка, распознавание изображений, принятие решений, прогнозирование и адаптация. Подобное определение требует философской проясненности. Мы имеем дело не с простой машинной автоматизацией, а с системой, которая участвует в действиях, имеющих интенциональную структуру – направленных на смысл, цель и ценности. Тем самым ИИ становится претендентом на область, традиционно принадлежавшую человеческому разуму.

Влияние ИИ на экономику и труд очевидно. Автоматизация, ускорение бизнес-процессов, роботизация производства – все это трансформирует рабочую среду и формы занятости. Тем не менее, за этой трансформацией скрыт более глубокий вопрос: если труд – как утверждали Маркс, Арендт, Сартр – является способом самореализации и источником человеческой идентичности, то что происходит, когда машины вытесняют человека из этого пространства? Не теряем ли мы саму категорию субъекта как активного деятеля?

Однако важно отметить, что с утратой одних профессий, появляются и новые – связанные с программированием, анализом данных, этическим аудитом ИИ и другими высокотехнологичными задачами. Это смещение требует гибкой, ориентированной на обучение и переосмысление социальной роли человека новой философии труда.

Одним из центральных вызовов, связанных с ИИ, является этическая дилемма. Классический пример – дилемма «трамвайной этики», которая впервые была сформулирована философом Филиппой Фут (1967), позже развита Джудит Джарвис Томсоном [1]. Суть вопроса: допустимо ли пожертвовать одним человеком, чтобы спасти пятерых? Если исходить из утилитаризма (Дж. Бентам, Дж. С. Милль), то нужно минимизировать вред, переключая стрелку и в итоге будет только одна жертва. Если же исходить из этики добродетели (Аристотель) решение должно приниматься в зависимости от контекста, характера и добродетели субъекта. В общем, механизм принятия решения, в одном случае может быть справедливым, в другом может быть этически неприемлемым. Вопросы, связанные с автономией и ответственностью ИИ, требуют от философов новых подходов к моральной теории.

Если мы доверяем такие решения ИИ, встает такой вопрос - может ли машина быть моральным агентом? По Канту, мораль предполагает автономию воли, т.е. способность подчиняться закону, выбранному как универсальному. Если ИИ действует по заранее заданному алгоритму, то он, возможно, не может быть носителем моральной ответственности.

В научной фантастике, от «Метрополиса» до современных киновселенных, часто возникает образ ИИ, который выходит из-под контроля создателя. Эти сюжеты выполняет не только развлекательную, но и философскую функцию, они моделируют проблему автономии творения. Если беспилотный автомобиль попадет в аварию или медицинский алгоритм ошибется в диагнозе — кто виноват? Можно ли научить ИИ отличать «хорошее» от «плохого», если понятие «хорошего» и «плохого» в разных культурах и ситуациях сильно варьируется? Не менее важный вопрос, а что, если ИИ станет сравним с человеческим разумом (или превзойдет его) — сможет ли человек контролировать такого «субъекта»? Несет ли человек моральную ответственность за то, что создал, но уже не контролирует? Может ли нравственность без запрограммирована, или она требует субъективного понимания?

Современные ИИ-системы демонстрируют впечатляющие способности к обучению и адаптации. В здравоохранении они распознают опухоли на ранних стадиях, обрабатывая миллионы изображений с поразительной точностью [2].

Однако философский вопрос здесь не в эффективности, а в природе понимания. Дж. Серль в своей гипотезе «китайской комнаты» утверждает, что даже если машина идеально имитирует понимание китайского языка, она не осознает смысла сказанного. Она лишь манипулирует символами.

Следовательно, можно задаться вопросом: обладает ли ИИ сознанием, или он лишь симулирует его? Для философов, таких как Даниэль Деннетт, сознание – это «высшая форма вычислительного процесса», и потому ИИ может быть сознательным в функциональном смысле. Другие, как Дж. Серль, утверждают, что без субъективного опыта сознание невозможно.

С развитием ИИ встает еще один важный философский вопрос, а что, если реальность, в которой мы живем – симулирована. Если машины способны имитировать или даже моделировать человеческие чувства и переживания, то как это влияет на наше понимание реальности? В частности, идеи о том, что ИИ может создать симуляцию мира, которая будет неотличима от реальности, напрямую связаны с философскими традициями, такими как теория симуляции, предложенная Н. Бостромом. Согласно этой теории, возможно, что наш мир — это всего лишь высокотехнологичная симуляция, созданная с помощью ИИ [3]. Такие концепции поднимают вопросы о том, как мы можем понимать «реальность», если вся наша жизнь и восприятие могут быть результатом работы сложных алгоритмов. В метафизике ИИ часто упоминается в контексте философских проблем о возможном существовании параллельных миров, искусственных реальностей и гипотезах о «моделировании» всего существующего. Подобные идеи отсылают к платоновской аллегории пещеры, к картезианским сомнениям, и даже к буддийскому учению о майе – иллюзии бытия. Если машина может создать симулированную реальность, неотличимую от «настоящей», то теряется устойчивость между бытием и кажимостью.

Какое значение тогда имеет истина? Как мы можем доказать, что наше сознание не встроено в искусственно созданную среду? Эти вопросы возвращают философию к самому ее истоку – к исследованию основания реальности.

Вместе с техническим развитием ИИ возникает вопрос о будущем самого человека. В эпоху быстрого технологического прогресса многие философы начали рассматривать концепцию «технофутуризма» — идеи о том, что человечество может перейти на новый уровень существования через симбиоз с технологиями, включая ИИ. В этом контексте возникает понятие «трансгуманизма», которое утверждает, что с помощью технологий можно значительно улучшить человеческие способности, как физические, так и когнитивные. Но вместе с этим появляются опасения, связанные с возможностью утраты человечности. Какие моральные и философские основания будут у нас для сохранения человеческого достоинства и свободы в условиях, когда ИИ и другие технологии могут изменить основные аспекты человеческой природы? Некоторые философы, такие как Жан-Бенедикт Леви, предупреждают о том, что чрезмерная зависимость от ИИ может привести к разрушению экзистенциальной автономии человека и упрощению жизни, что лишит людей возможности размышлять о своем месте в мире.

В эпоху «постчеловеческого» философия сталкивается с необходимостью реантропологизации – переосмысления человека не только как биологического вида, но как существа, способного к нравственному выбору, творчеству и самотрансценденции.

ИИ – не только технологическая инновация, а философский вызов. Он ставит под сомнение вековые представления о субъекте, разуме, морали, реальности и будущем. Перед философией встает задача переосмысления того, что значит быть человеком.

Сегодня ИИ можно назвать зеркалом, в которое человечество смотрит с тревогой и восхищением. Оно показывает нам не только иные формы интеллекта, сколько глубины нашей собственной природы. Границы между творцом и творением, инструментом и субъектом, программой и душой становятся все менее четкими. Философия ИИ – это философия человека, обратившегося к собственному предельному вопросу, а именно, что значит быть мыслящим существом – в мире, где мыслит не только он?

#### Список использованной литературы:

1. Эдмондс Д. Убили бы вы толстяка? Задача о вагонетке: что такое хорошо и что такое плохо? М. : Изд-во Института Гайдара, 2016
2. Казахстанские студенты заявили, что разработали нейросеть, которая может выявить рак за 7 минут. IT-компания NUR.KZ [Электронный ресурс]: Режим доступа: <https://www.nur.kz/technologies/software/2059166-kazahstanskije-studenty-zayavili-chto-razrabotali-neyroset-kotoraya-mozhet-vyyavit-rak-za-7-minut/>
3. Nick Bostrom. Are We Living in a Computer Simulation? // The Philosophical Quarterly: journal. — Oxford University Press, 2003. — Vol. 53, no. 211. — P. 243—255.