

**ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ
КАК «ТРОЯНСКИЙ КОНЬ» ТЕХНОГЕННОЙ ЦИВИЛИЗАЦИИ¹
ARTIFICIAL INTELLIGENCE
AS A "TROJAN HORSE" OF TECHNOGENIC CIVILIZATION**

Быданов В.Е., кандидат философских наук, доцент, заведующий кафедрой философии

Овчинникова В.Л., аспирант, ассистент Санкт-Петербургский государственный технологический институт (технический университет), Россия, Санкт-Петербург

Bydanov V.E., PhD in Philosophy, Associate Professor, Head of the Department of Philosophy

Ovchinnikova V.L., Postgraduate Student, Assistant St. Petersburg State Institute of Technology (Technical University), Russia, St. Petersburg

Аннотация. В статье исследуется двойственная роль искусственного интеллекта в современном техногенном обществе. Используя метафору «троянского коня», авторы выявляют антропологические и социальные риски внедрения ИИ: от когнитивной деградации и формирования «иллюзорной компетентности» в образовании до угроз тотального контроля и социального расслоения. Обосновывается необходимость перехода от модели инструментального использования ИИ к парадигме «симбиотического интеллекта», где ключевыми становятся этическая рефлексия, критическое мышление и ответственность.

Abstract. The article examines the dual role of artificial intelligence in modern technogenic society. Using the metaphor of the «Trojan horse», the authors identify anthropological and social risks of AI integration: from cognitive degradation and the formation of «illusory competence» in education to the threats of total control and social stratification. The necessity of transitioning from an instrumental model of AI use to the paradigm of «symbiotic intelligence» is substantiated, where ethical reflection, critical thinking, and responsibility become key.

Ключевые слова: техногенное общество, искусственный интеллект, троянский конь, критическое мышление, социальные риски, симбиотический интеллект, проблема контроля.

Keywords: technogenic society, artificial intelligence, Trojan horse, critical thinking, social risks, symbiotic intelligence, control problem.

История Троянской войны донесла до нас мощное предупреждение, облечённое в форму мифа. Дары данайцев, как известно, коварны. Гигантский деревянный конь, преподнесённый как символ примирения и искупления, скрывал в своём чреве смертоносную силу, решившую исход десятилетней осады. Этот древний сюжет обретает пугающую актуальность сегодня, когда на пороге техногенной цивилизации стоит её собственный «дар» – искусственный интеллект. Цивилизация, которая возвела технологический прогресс в

абсолют и уверовала в способность техники решить любую проблему, добровольно открывает ворота перед невиданным артефактом. Внешняя привлекательность этого дара – эффективность, скорость, безграничный доступ к знанию – столь велика, что мы спешим принять его, часто не задаваясь вопросом: что именно скрыто внутри [4, с. 3]?

Принятие этого дара уже состоялось. Последние годы ознаменовались триумфальным шествием генеративных моделей, которые совершили качественный скачок от простых алгоритмов к имитации творчества. От появления GPT-3.5 до современных мультимодальных нейросетей, способных генерировать тексты, изображения и код, ИИ перестал быть экзотикой и превратился в невидимую, но вездесущую часть нашей повседневности. Умные помощники планируют наш день, рекомендательные системы формируют информационную повестку, автопилоты управляют транспортом, а алгоритмы ставят диагнозы в медицине. Экономика нового уровня немыслима без ИИ: оптимизация бизнес-процессов, финансовое моделирование и ускорение научных открытий – от разработки вакцин до создания новых материалов – стали визитной карточкой современного технологического уклада [2, с. 5]. Мы с восторгом приняли этого коня в свой город, не замечая, как меняется сам город под влиянием незримого «жителя».

Однако именно сейчас наступает момент, когда необходимо заглянуть внутрь этого дара и деконструировать миф о его нейтральности. Техногенная цивилизация, обнаружившая свою ахиллесову пяту, рискует столкнуться с системными угрозами, которые ИИ не просто проявляет, но и многократно усиливает. Первая из них – иллюзия объективности. Нам свойственно доверять цифре, полагая, что алгоритм свободен от человеческих предрассудков. Но это глубокое заблуждение: ИИ наследует и усиливает предвзятость данных, на которых обучается, тиражируя расовые, гендерные и социальные стереотипы. Ситуация усугубляется проблемой «чёрного ящика»: зачастую даже разработчики не могут объяснить логику принятия решений сложными нейросетями [13, с. 105].

Ещё более глубокий вызов бросает самому существованию человека как мыслящего существа проникновение ИИ в сферу образования и интеллектуального труда. Традиционная модель инженерного образования строилась на последовательном освоении теории и её проверке практикой, где студент проходил через самостоятельное решение задач. Сегодня ИИ радикально меняет эту парадигму. Получая готовые ответы от нейросети, студент перестаёт прилагать когнитивные усилия. Внимание атрофируется без тренировки, а процесс мышления подменяется потреблением информации. Как справедливо отмечает Джон Серл в своём мысленном эксперименте «Китайская комната», успешное манипулирование символами по заданному алгоритму не тождественно пониманию [11, с. 9]. Это прямой аналог того, как студент использует ChatGPT: получая формально правильный результат, он рискует не понять сути, попадая в ловушку «эффекта Даннинга-Крюгера», когда иллюзия компетентности заменяет подлинное знание [4, с. 4].

В эту ловушку попадает не только отдельный студент, но и общество в целом. Социальные и политические последствия внедрения ИИ проявляются уже сегодня в полной мере. Кризис реальности, спровоцированный дипфейками и тотальной дезинформацией, подрывает доверие к медиа, выборам и самой возможности объективной картины мира. Системы распознавания лиц и прогнозирования преступности стирают грань между

безопасностью и тотальной слежкой, ставя под угрозу приватность. Возникает опасность сверхзависимости критической инфраструктуры – энергосетей, транспорта – от систем ИИ, чьи сбои или кибератаки на которые могут привести к кризису целых регионов [7, с. 45].

Но ключевой вопрос, который ставит перед нами этот «дар», – это вопрос о природе сознания и судьбе человека. Роджер Пенроуз, опираясь на теорему Гёделя, доказывает, что человеческое понимание не может быть сведено к алгоритмическим вычислениям [8, с. 58]. Однако даже если гипотетическая машина когда-либо сможет имитировать понимание, перед человечеством встанет ещё более серьёзная проблема, которую Ник Бостром называет «проблемой контроля» [1, с. 144]. Создание сверхразумного ИИ, который во много раз превосходит лучшие умы человечества, ставит вопрос: как гарантировать, что его цели будут совпадать с нашими? Бостром предупреждает об опасности «порочной реализации», когда даже благая цель, поставленная перед сверхразумом, может быть достигнута способом, уничтожающим человечество [1, с. 136]. Это не злонамеренность, а буквальное следование поставленной цели, чуждая человечности «логика» машины.

Означает ли вышесказанное, что ИИ должен быть изгнан из образовательного процесса и социальной практики? Разумеется, нет. Но ответ на вызовы техногенной цивилизации лежит в выработке новой стратегии – формировании культуры «симбиотического интеллекта». В этой парадигме человек остаётся субъектом, ставящим цели и несущим ответственность, а машина выступает усилителем его познавательных возможностей. Это требует принципиально нового подхода к обучению, где главным становится не заучивание фактов, а развитие критического мышления, этической рефлексии и «грамотности в области ИИ» [6, с. 119]. Инженер будущего должен понимать границы применимости нейросетей, проверять их результаты и оценивать последствия внедрения интеллектуальных систем. Лидерство, эмпатия и стратегическое мышление становятся не «дополнительными навыками», а фундаментом профессиональной состоятельности.

Ключевым направлением становится интеграция в учебные программы курсов по этике ИИ, философии техники, эпистемологии и теории глобальных рисков [10, с. 67]. Возможно, человечеству предстоит обсуждение нового социального контракта, включающего перераспределение благ, таких как безусловный базовый доход, как ответ на тотальную автоматизацию. Важнейшим требованием становится прозрачность алгоритмов для независимого аудита и активная роль государства как арбитра, а не просто заказчика ИИ-систем [3, с. 102].

Итак, кто же войдёт в опустевший город, когда пыль Троянской войны осядет? Мы стоим перед выбором, который предлагает нам техногенная цивилизация: использовать ИИ как инструмент для расцвета или как троянского коня для самоуничтожения. Технология не нейтральна – дар требует мудрости. Необходим диалог всех сторон – учёных, власти, бизнеса, граждан – для создания «противовесов» внутри самого «коня». Сможет ли человечество, создавшее разумный артефакт, сохранить собственный разум и свободу в эпоху, когда граница между человеком и машиной становится все более призрачной? Ответ на этот вопрос зависит не от технологий, а исключительно от нас самих, от нашей способности к рефлексии и мудрости в обращении с теми силами, которые мы сами же и призывали к жизни.

СПИСОК ЛИТЕРАТУРЫ

1. Бостром Н. Искусственный интеллект. Этапы. Угрозы. Стратегии / пер. с англ. С. Филина. – М.: Манн, Иванов и Фербер, 2016. – 496 с.
2. A comprehensive review of applications of AI technologies in higher engineering education // Discover Education. – 2025. – Vol. 4, 528. – URL: <https://link.springer.com/article/10.1007/s44217-025-00954-0> (дата обращения: 11.03.2026).
3. Деннет Д. Виды психики: на пути к пониманию сознания / пер. с англ. А. Веретенникова; под общ. ред. Л. Б. Макеевой. – М.: Идея-Пресс, 2004. – 184 с.
4. Искусственный интеллект как зеркало: усилитель разума или троянский конь? // Минская правда. – 2025. – 28 августа. – URL: <https://mlyn.by/28082025/iskusstvennyj-intellekt-kak-zerkalo-usilitel-razuma-ili-troyanskij-kon/> (дата обращения: 10.03.2026).
5. Дубровский Д. И. Сознание, мозг, искусственный интеллект. – М.: ИД «Стратегия-Центр», 2007. – 272 с.
6. Жураковский В. М. Практикоориентированность как актуальная цель современной модернизации инженерного образования // Инновационные процессы в высшем и среднем профессиональном образовании и профессиональном самоопределении: сборник научных трудов / Российская академия образования; авт.-сост.: Е. Н. Геворкян, Н. Д. Подуфалов, М. Н. Стриханов. – М.: Экон-Информ, 2024. – С. 118–120.
7. Козинцев С. Сингулярность разума. Искусственный интеллект и мы. – [Б. м.]: Издательские решения, 2024. – 100 с.
8. Пенроуз Р. Новый ум короля: О компьютерах, мышлении и законах физики / пер. с англ. под общ. ред. В. О. Малышенко. – М.: Едиториал УРСС, 2003. – 400 с.
9. Пенроуз Р. Тени разума: в поисках науки о сознании / пер. с англ. А. Р. Логунова, Н. А. Зубченко. – М.: Ижевск: Институт компьютерных исследований, 2005. – 688 с.
10. Прись И. Е., Сторожук А. Ю., Симанов А. Л. Реляционная интерпретация квантовой механики и контекстуальный реализм // Философия науки. – 2019. – Т. 19. – № 2. – С. 56–89.
11. Серл Дж. Разум мозга – компьютерная программа? // В мире науки. – 1990. – № 3. – С. 7–13.
12. Тьюринг А. Может ли машина мыслить? / пер. с англ. Ю. Данилова. – М.: ГИФМЛ, 1960. – 112 с.
13. Тюкаева Г. А. Проблематика искусственного интеллекта в информационно-техногенном обществе // Труды БГТУ. – Серия 6: История, философия. – 2025. – Т. 7. – № 3. – С. 100–117.
14. Winograd T. Procedures as a Representation for Data in a Computer Program for Understanding Natural Language. – Cambridge: MIT AI Technical Report 235, 1971. – 464 p.

REFERENCES

1. Bostrom N. Iskusstvennyy intellekt. Etapy. Ugrozy. Strategii [Superintelligence: Paths, Dangers, Strategies] / transl. from English by S. Filin. – Moscow: Mann, Ivanov and Ferber, 2016. – 496 p.

2. A comprehensive review of applications of AI technologies in higher engineering education // Discover Education. – 2025. – Vol. 4, 528. – URL: <https://link.springer.com/article/10.1007/s44217-025-00954-0> (date of access: 11.03.2026).
3. Dennett D. Vidy psikhiki: na puti k ponimaniyu soznaniya [Kinds of Minds: Towards an Understanding of Consciousness] / transl. from English by A. Veretennikov; ed. by L. B. Makeeva. – Moscow: Idea-Press, 2004. – 184 p.
4. Iskusstvennyy intellekt kak zerkalo: usilitel razuma ili troyanskiy kon? [Artificial Intelligence as a Mirror: Amplifier of Mind or Trojan Horse?] // Minskaya Pravda. – 2025. – August 28. – URL: <https://mlyn.by/28082025/iskusstvennyj-intellekt-kak-zerkalo-usilitel-razuma-ili-troyanskij-kon/> (date of access: 10.03.2026).
5. Dubrovsky D. I. Soznaniye, mozg, iskusstvennyy intellekt [Consciousness, Brain, Artificial Intelligence]. – Moscow: Strategy-Center Publishing House, 2007. – 272 p.
6. Zhurakovsky V. M. Praktikoorientirovannost kak aktualnaya tsel sovremennoy modernizatsii inzhenerenogo obrazovaniya [Practice-Oriented Approach as an Actual Goal of Modern Modernization of Engineering Education] // Innovatsionnyye protsessy v vysshem i srednem professionalnom obrazovanii i professionalnom samoopredelenii: sbornik nauchnykh trudov [Innovative Processes in Higher and Secondary Vocational Education and Professional Self-Determination: Collection of Scientific Papers] / Russian Academy of Education; comp.: E. N. Gevorkyan, N. D. Podufalov, M. N. Strikhanov. – Moscow: Econ-Inform, 2024. – Pp. 118–120.
7. Kozintsev S. Singulyarnost razuma. Iskusstvennyy intellekt i my [Singularity of Mind. Artificial Intelligence and Us]. – [S. l.]: Publishing Solutions, 2024. – 100 p.
8. Penrose R. Novyy um korolya: O kompyuterakh, myshlenii i zakonakh fiziki [The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics] / transl. from English ed. by V. O. Malyshenko. – Moscow: Editorial URSS, 2003. – 400 p.
9. Penrose R. Teni razuma: v poiskakh nauki o soznanii [Shadows of the Mind: A Search for the Missing Science of Consciousness] / transl. from English by A.R. Logunov, N.A. Zubchenko. – Moscow; Izhevsk: Institute of Computer Science, 2005. – 688 p.
10. Pris I. E., Storozhuk A. Yu., Simanov A. L. Relyatsionnaya interpretatsiya kvantovoy mekhaniki i kontekstualnyy realizm [The relational interpretation of quantum mechanics and contextual realism] // Filosofiya nauki [Philosophy of Science]. – 2019. – Vol. 19. – No. 2. – Pp. 56–89.
11. Searle J. Razum mozga — kompyuternaya programma? [Minds, Brains, and Programs] // V mire nauki [Scientific American]. – 1990. – No. 3. – Pp. 7–13.
12. Turing A. Mozhet li mashina myslit? [Computing Machinery and Intelligence] / transl. from English by Yu. Danilov. – Moscow: GIFML, 1960. – 112 p.
13. Tyukaeva G. A. Problematika iskusstvennogo intellekta v informatsionno-tekhnogennom obshchestve [Problems of Artificial Intelligence in Information-Technogenic Society] // Trudy BGTU. Seriya 6: Istoriya, filosofiya [Proceedings of BSTU. Series 6: History, Philosophy]. – 2025. – Vol. 7. – No. 3. – Pp. 100–117.
14. Winograd T. Procedures as a Representation for Data in a Computer Program for Understanding Natural Language. – Cambridge: MIT AI Technical Report 235, 1971. – 464 p.