

2. Sokolov V.M. Sociologiya morali – real'naya ili gipoteticheskaya // Sociologicheskie issledovaniya. 2004. № 8.
3. Kosharnaya G.B., Afanas'eva Yu.L. Cennostnye orientacii sovremennoj rossijskoj molodezhi // Izvestiya vysshih uchebnyh zavedenij. Povolzhskij region. Obshchestvennye nauki. Sociologiya. 2008. № 4.
4. Rusakova D.M. Cennostnye orientacii sovremennoj molodezhi // Mir nauki. Sociologiya, filologiya, kul'turologiya. 2018. №1. t. 9.
5. Shlygin S.P., Obishchenko A.V. Formirovanie cennostnyh orientacij sovremennoj rossijskoj molodezhi // Aktual'nye issledovaniya. 2020. №18 (21).
6. Rudkevich E.Yu. Sistema cennostej obshchestva: strukturnyj analiz // Vlast'. 2007. № 1.
7. Gryaznova E.V., Afanas'ev S.V. O sootnoshenii ponyatij «cennosti», «duhovnye cennosti» i «kul'turnye cennosti» // Innovacionnaya ekonomika: perspektivy razvitiya i sovershenstvovaniya. 2019. № 5.
8. Simonenko N. I. Ekologicheskaya kul'tura v sovremenom sociokul'turnom diskurse: k opredeleniyu ontologicheskoy sushchnosti ekologicheskoy kul'tury // Vestnik Chelyabinskogo gosudarstvennogo universiteta. 2009. № 29.
9. Gurko T. A. Novye semejnye formy: tendencii rasprostraneniya i ponyatiya // Sociologicheskie issledovaniya. 2017. № 11.
10. Shchemeleva I. I. Social'naya aktivnost' studencheskoj molodezhi: faktornyj i klasternyj analiz // Sociologicheskie issledovaniya. 2019. № 4.

**РУБЦОВ АЛЕКСАНДР ГЕННАДЬЕВИЧ** - ассистент кафедры истории и философии, РЭУ им. Г.В. Плеханова (Rubchic65@yandex.ru).

**RUBTSOV, ALEXANDER G.** - Assistant of the Department of History and Philosophy, Plekhanov Russian University (Rubchic65@yandex.ru).

УДК 172

DOI: 10.24412/2411-2275-2023-4-263-267

**ИВЛЕВ Д.В.**

## **ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ И ПРОБЛЕМЫ ЭТИКИ**

**Ключевые слова:** этика, искусственный интеллект, социальная философия, нейросети, умный город, социальные риски

В статье дается обзор этических вопросов, восходящих к проблематике взаимоотношений «искусственный интеллект- общество» и «искусственный интеллект – человек», через призму социально-философской проблематики, эстетики и практической философии. Отмечается разнообразие точек зрения исследователей по вопросу об оценке социальных последствий повсеместного внедрения искусственного интеллекта в социальную жизнь, анализируются возможные риски этого процесса. Особое внимание уделяется вопросам этики внедрения социальных роботов, беспилотного транспорта, применения искусственного интеллекта в творчестве, социально-психологическим аспектам взаимодействия человека и искусственного интеллекта.

**IVLEV, D.V.**

### **ARTIFICIAL INTELLIGENCE AND ETHICAL ISSUES**

**Key words:** ethics, artificial intelligence, social philosophy, neural networks, smart city, social risks

This article provides an overview of ethical issues related to the problems of the relationship "artificial intelligence- society" and "artificial intelligence – man" through the prism of socio-philosophical issues, aesthetics and practical philosophy. The diversity of researchers' points of view on the assessment of the social consequences of the widespread introduction of artificial intelligence into social life is noted, and the possible risks of this process are analyzed. Special attention is paid to the ethics of the introduction of social robots, unmanned vehicles, the use of artificial intelligence in creativity, socio-psychological aspects of human interaction and artificial intelligence.

Последние десять лет стали ключевыми в области развития искусственного интеллекта. От времени первых голосовых помощников, представленных широкой аудитории в начале 2010-х годов до триумфа ChatGPT в 2022 году прошло не многим более десяти лет, но уже сейчас очевидно, что искусственный интеллект играет все более и более важную роль, проникая во все новые и новые сферы профессиональной деятельности и личного пространства индивида.

Современные ученые активно исследуют феномен искусственного интеллекта, как с точки зрения информатики, науковедения, так и в социально-гуманитарных науках.

Э.М. Пройдаков в статье «Современное состояние искусственного интеллекта» представил краткую историю работ в области искусственного интеллекта, охарактеризовал направления разработок в этой области, привел общий обзор современного состояния исследований систем искусственного интеллекта, а также проанализировал возможные области его применения и проблемы, с которыми человечество может столкнуться в условиях его активного использования. Он разделил системы искусственного интеллекта на два уровня: общий, или универсальный, в принципах деятельности приближающийся к копированию деятельности человеческого мозга, и поэтому способный к обучению и многозадачности, а также слабый, или прикладной, который, несмотря на масштабы и скорость обрабатываемой информации, выполняет задачи лишь в конкретной области, и его целью является помощь в повышении производительности определенных видов работы.

Он проследил историю возникновения искусственного интеллекта, начиная с 50-х годов XX века, в первых шахматных программах и программах машинного перевода, а также, немного позднее, в программах для доказательства математических теорем и решения задач. Последующие этапы развития искусственного интеллекта были связаны с работами в области логического

программирования и генетического программирования, созданием программных агентов. Современный этап развития искусственного интеллекта характеризуется эпохой технологии искусственных нейронных сетей, которые имитируют работу нейронов в мозговой деятельности живых существ.

В статье также были классифицированы риски широкого применения искусственного интеллекта в социальной и общественной жизни. К этим рискам относятся такие, как переход к полной зависимости от компьютеров, в том числе, в области управления и принятия решений, неконтролируемое появление новых потенциально опасных технологий, непредсказуемые последствия искусственного интеллекта в производстве вооружений, проявление ошибок, допущенных в период машинного обучения в системах, управляемых искусственным интеллектом, а также социальные риски: «в первую очередь прогнозируемое исчезновение целого ряда профессий (например, водителей, младшего медперсонала и даже журналистов), разобщение людей, а возможно, даже потерю естественных навыков человека. Развитие техники постоянно идет по пути исключения человека из выполняемых им производственных действий. Процессы эти сложные, потому что с исчезновением старых профессий появляются новые, связанные с новой цифровой экономикой, внедрением и поддержанием новой инфраструктуры. Однако в целом потребности в рабочей силе должны сильно уменьшиться, поскольку машины могут заменить человека в большинстве видов активной (не творческой) деятельности. Интеллектуализация, цифровизация и роботизация могут привести как к появлению в обществе социальной напряженности, так и к внедрению мер для замедления указанных процессов» [1, с. 149].

В центре внимания некоторых исследователей находятся только лишь возможные позитивные последствия широкого внедрения искусственного интеллекта в жизнь общества. Так, М.А. Ижунинов в статье «Обзор современных нейронных сетей и их интеграция в жизнь человечества» оптимистично оценивает разнообразие имплементации нейросетей в экономику и общество и рассматривает как положительный момент, что искусственный интеллект, благодаря скорости обработки информации, освободит человека от необходимости владеть разнообразной информацией, сосредоточившись на навыках управления компьютером и знаниями в области своей деятельности «В любом случае искусственная нейронная сеть имеет больше преимуществ, чем недостатков, самое главное, держать эту технологию под контролем и не допустить, чтобы она попала не в те руки» [2, с. 19].

Сам по себе тезис о желательном избавлении от работы с «ненужной информацией» и сосредоточении исключительно на узкоспециализированной информации определенной сферы деятельности является неоднозначным и объективно имеет своим следствием обеднение интеллектуального мира человека, и, в дальнейшем, слабое владение компетенциями обработки и анализа общественной, культурной и социально-политической информации. В целом, изучая последствия неизбежной имплементации искусственного интеллекта в жизнь современного социума, большинство исследователей отмечают необходимость их тщательного разностороннего анализа, при условии потенциальных рисков. Так, Т.В. Семина считает, что «самый большой риск в эпоху искусственного интеллекта – недостаточное осознание людьми опасностей применения данной технологии. Нельзя позволить, чтобы по причине удобной технической иллюзии люди утратили свое субъективное сознание и автоматически потеряли свою «истинную свободу» как базис человеческого существа» [3, с. 176]. А.С.Вантяева рассматривает вопрос распространения искусственного интеллекта с точки зрения принципа гуманизма: «происходит активная технократизация общественных сфер жизнедеятельности, что несет ряд потенциальных угроз и рисков. Например, риск дегуманизации, так как с появлением умных машин уменьшается значимость саморазвития человека. Возникает феномен иррациональной рациональности, который приводит к дегуманизации человеческого разума, что впоследствии может привести к угрозе человечеству как более слабому виду» [4, с. 69].

Е.Е. Хабирова, на основе проведенных социологических исследований отношения российского населения к возможностям применения искусственного интеллекта, делает вывод о том, что «по мере возрастания степени самостоятельности и ответственности принятия решений на основе искусственного интеллекта доверие населения снижалось» [5, с. 266] (если 66 % опрошенных было бы комфортно иметь домашнего робота-помощника, то участвовать в

совместной трудовой деятельности с роботами были готовы 44 % респондентов, а к использованию роботов в хирургии готово только 14 %).

Пройдя путь от инструмента, автоматизирующего процесс языкового перевода текста и логистического учета, искусственный интеллект постепенно становится полноценным важным фактором бытия современного человека, который не только помогает решать те или иные конкретные поставленные перед ним задачи, но и участвует в формировании картины мира, эстетических вкусов, а в будущем вполне возможно будет генерировать аксиологические императивы человека и общества.

Такое понимание текущего состояния взаимоотношений «человек-искусственный интеллект» формирует проблематику уже не с позиции технологии, а с позиций этики: какие потенциальные угрозы несет в себе слияние человека и нейросетевых алгоритмов в обыденной жизни, следует ли законодательно ограничить внедрение искусственного интеллекта в ткань общества или проблема преувеличена? На эти и другие этические вопросы современные исследователи выдвигают ответы, исходя из разных точек зрения, а также предлагают возможные пути их решения.

Важно отметить, что в этических исследованиях просматриваются два уровня проблематики: во-первых, это система «искусственный интеллект - общество», во-вторых, система взаимоотношений «искусственный интеллект — человек». К вопросам первой системы прежде всего относятся перспективы развития среды умного города, внедрение среды беспилотного транспорта, роботизация; в рамках системы отношений человека и искусственного интеллекта поднимаются вопросы общения и психологического состояния дружбы с нейронными сетями, а также получение специального сгенерированного развлекательного контента от нейросетей. (развитие среды умного города, внедрение беспилотного транспорта, роботизация) (общение и дружба с искусственным интеллектом, получение специального сгенерированного развлекательного контента от нейросетей). В первом случае поднимаются вопросы этики в прочной связи с проблемами социальной философии, а во втором — этический аспект в рамках антропологии: каким будет человек, формирующийся под влиянием искусственного интеллекта, в альтернативах добра и зла?

Умный город в современном понимании - это интернет вещей, цифровизация услуг и искусственный интеллект, работающий на благо общества. В таком пространстве каждый человек добровольно передает большое количество информации о себе городским службам и бизнесу. Естественным образом встает вопрос этики обработки и безопасного хранения данных, например, информации о распознавании лиц и передаче ее третьим лицам (бизнесу). С одной стороны, такая технология позволяет поддерживать высокий уровень безопасности в общественных пространствах, но с другой вызывает у исследователей озабоченность по поводу личного пространства и свободы отдельно взятого жителя города [6, с. 4].

В России в умных городах и наукоградах в настоящее время ведется апробация технологий автоматизированных роботов-доставщиков, тестирование беспилотного такси и грузоперевозок [7] для потенциального массового внедрения в обозримом будущем. В целом направление беспилотного транспорта выстраивается с позиции установки безопасности человека в качестве главного приоритета, и по предположению исследователей проблемы, опыт получаемый из данного направления, должен лечь в основу будущей этической карты автономных роботов, которые будут непосредственными участниками социального взаимодействия, индустрии развлечений или медицины [8, р. 96]. При этом полностью перенести этический опыт, полученный из области внедрения беспилотного транспорта, также не представляется возможным, так как, к примеру, безопасность информации, передаваемой клиентом беспилотного такси, гораздо менее критична, чем то качество и объем информации, что будет получать «социальный» робот.

Рассматривая вопрос этики применения социальных роботов, необходимо отметить, что внедрение такого типа автоматических систем широко обсуждается в сфере ухода за пожилыми людьми. Исследователями этой области раскрывается важная этическая проблематика: распространение современных человекоподобных роботов сказывается на качестве оказываемой помощи, несмотря на те преимущества, которые появляются в процессе их внедрения. Роботы пока не могут обеспечить тот же уровень внимания к пациентам, что и люди-специалисты,

осуществляющие уход; в исследовании отмечается важность объединений усилий зрителей за пожилыми людьми, социальных роботов и их разработчиков для разрешения потенциальных этических проблем; а также предлагается этическая дорожная карта в помощь будущим создателям подобных роботов [9, p. 1640].

Рассуждая в менее глобальном масштабе, на уровне взаимоотношений «искусственный интеллект – индивидуум», исследователи спорят о том, насколько работы нейросетевых алгоритмов можно считать искусством/творчеством [10]. Но для конечного, массового потребителя творческой продукции вопросы философской эстетики не так принципиальны, и при появлении возможности приобретения качественного специально сгенерированного художественного материала под свои вкусы, пользователь рационально выберет именно такой вариант. Пускай в современном мире и растет объем рынка креаторов - креативной экономики [11, с. 99], искусственный интеллект, благодаря своей скорости производства контента, гибкости и доступности для конечного пользователя может стать не только вспомогательным инструментом для авторов-креаторов, но и полноценным конкурентом в их творческой деятельности. Что опять же подводит исследователей к вопросу: насколько этически правильно оставлять область искусства на откуп искусственному интеллекту?

При этом потенциал искусственного интеллекта в создании творческого материала поистине колоссален. Музыка, живопись, литература, а в перспективе кинематограф, - сложно представить сферу искусства, где нейросети не оставят свой след. Так, в архитектуре и дизайне помещений прослеживается большой задел для использования генеративных моделей, работа которых может привести к созданию новых архитектурных стилей [12]. Создавая визуальный нарратив вокруг человека, нейросети будут оказывать существенное влияние на чувство вкуса зрителя, и принятие решений о том, что красиво и безобразно останется в серой зоне: будет ли у человека при таких условиях свобода эстетического выбора?

На более глубоком уровне взаимодействия человека с ИИ исследователями прослеживается дружеская связь индивидуума и машины, в которой человек начинает воспринимать виртуального собеседника в качестве воображаемого друга или как отражения себя в зеркале [13, p. 6]. К такому «другу», к которому можно обратиться, когда рядом нет собеседников из числа существующих в реальном мире. То, как повлияет такой симбиоз на бытие человека в перспективе — открытый вопрос. Возможно, социальная интеракция с живым собеседником все равно останется в приоритете для большинства людей, а искусственный собеседник так и останется всего лишь голосовым помощником, а не полноценным другом.

В заключение следует отметить, что те области, в которых искусственный интеллект уже проявил себя как полноправный участник отношений с человеком, вопросы этики возникают естественным образом, ведь искусственный интеллект прямо или косвенно заменяет человеческую деятельность на свою собственную — «искусственную». Тем самым все те вопросы этики, которые задавали философы и общественные деятели по отношению к человеку и обществу, в настоящее время справедливо адресовать уже в отношении искусственного интеллекта и результатов его деятельности. Пускай пока искусственный интеллект и понимается больше как инструмент, а не как полноправный участник, формирующий реальность, этическая оценка его внедрения в общественную жизнь, - это необходимость для понимания потенциальных угроз и возможных непредвиденных последствий.

### Литература и источники

1. Пройдаков Э.М. Современное состояние искусственного интеллекта // Научно-исследовательские исследования. 2018. № 1.
2. Ижунин М. А. Обзор современных нейронных сетей и их интеграция в жизнь человечества // Молодой ученый. 2020. № 4 (294).
3. Семина Т.В. Воздействие технологий искусственного интеллекта на социальные отношения // Социология. 2022. № 3.
4. Ваняева А.С. Социальные риски внедрения технологий искусственного интеллекта // Теория и практика общественного развития. 2022. № 7.
5. Хабирова Е.Е. Неопределенность в отношении технологии искусственного интеллекта в российском обществе // Вестник Санкт-Петербургского университета. Социология. 2021. Т. 14. Вып. 3.
6. Tao M., Jiang R., Downs C. Ethics of Face Recognition in Smart Cities Toward Trustworthy AI / Big Data Privacy and Security in Smart Cities. Springer Cham, 2022.
7. Беспилотное такси в Иннополисе: цифры и факты // [URL] <https://yandex.ru/blog/company/bespilotnoe-taksi-v-innopolise-tsfy-i-fakty?ysclid=lp1k60ofq119428231> (дата обращения: 24.11.2023)

8. Dodig-Crnkovic G.D., Holstein T., Pelliccione P., Thavarasa J. Future Intelligent Autonomous Robots, Ethical by Design – Lessons Learned from Autonomous Cars Ethics / Proceedings of the 14th International Conference on Society and Information Technologies: ICSIT 2023. March 28 - 31, 2023.
9. Yuan S., Coghlan S., Lederman R., Waycott J. Ethical Design of Social Robots in Aged Care: A Literature Review Using an Ethics of Care Perspective // International Journal of Social Robotics. 2023. №15(9-10).
10. Coeckelbergh M. The Work of Art in the Age of AI Image Generation: Aesthetics and Human-Technology Relations as Process and Performance // Journal of Human-Technology Relations. 2023. Vol. 1. // [URL] <https://journals.open.tudelft.nl/jhtr/article/view/7025> (Дата доступа 30.11.2023).
11. Соловьева И.А. Креативная экономика - основа стратегического развития экономики // Лучшие практики развития креативных индустрий в регионах России. Материалы межрегиональной (с международным участием) научно-практической конференции / Под общей редакцией Ю.П. Соболевой, А.В. Амеличкина, Н.А. Илюхиной. Орел, 2023.
12. Ploennigs J., Berger M. AI art in architecture // AI in Civil Engineering. 2022. № 8.
13. Brandtzaeg P.B., Skjuve M.B., Følstad A. My AI Friend: How Users of a Social Chatbot Understand Their Human–AI Friendship // Human Communication Research. 2022.

### References and Sources

1. Projdakov E.M. Sovremennoe sostoyanie iskusstvennogo intellekta // Naukovedcheskie issledovaniya. 2018. № 1.
2. Izhuninov M. A. Obzor sovremennyh neyronnyh setej i ih integraciya v zhizn' chelovechestva // Molodoy uchenyj. 2020. № 4 (294).
3. Cemina T.V. Vozdejstvie tekhnologij iskusstvennogo intellekta na social'nye otnosheniya // Sociologiya. 2022. № 3.
4. Vantyaeva A.S. Social'nye riski vnedreniya tekhnologij iskusstvennogo intellekta // Teoriya i praktika obshchestvennogo razvitiya. 2022. № 7.
5. Habirova E.E. Neopredelennost' v otnoshenii tekhnologii iskusstvennogo intellekta v rossijskom obshchestve // Vestnik Sankt-Peterburgskogo universiteta. Sociologiya. 2021. T. 14. Vyp. 3.
6. Tao M., Jiang R., Downs C. Ethics of Face Recognition in Smart Cities Toward Trustworthy AI / Big Data Privacy and Security in Smart Cities. Springer Cham, 2022.
7. Bepilnoe taksi v Innpolise: cifry i fakty // [URL] <https://yandex.ru/blog/company/bepilnoe-taksi-v-innpolise-tsifry-i-fakty?ysclid=lp1k60ofq119428231> (data obrashcheniya: 24.11.2023)
8. Dodig-Crnkovic G.D., Holstein T., Pelliccione P., Thavarasa J. Future Intelligent Autonomous Robots, Ethical by Design – Lessons Learned from Autonomous Cars Ethics / Proceedings of the 14th International Conference on Society and Information Technologies: ICSIT 2023. March 28 - 31, 2023.
9. Yuan S., Coghlan S., Lederman R., Waycott J. Ethical Design of Social Robots in Aged Care: A Literature Review Using an Ethics of Care Perspective // International Journal of Social Robotics. 2023. №15(9-10).
10. Coeckelbergh M. The Work of Art in the Age of AI Image Generation: Aesthetics and Human-Technology Relations as Process and Performance // Journal of Human-Technology Relations. 2023. Vol. 1. // [URL] <https://journals.open.tudelft.nl/jhtr/article/view/7025> (Data dostupa 30.11.2023).
11. Solov'eva I.A. Kreativnaya ekonomika - osnova strategicheskogo razvitiya ekonomiki // Luchshie praktiki razvitiya kreativnyh industrij v regionah Rossii. Materialy mezhhregional'noj (s mezhdunarodnym uchastiem) nauchno-prakticheskoy konferencii / Pod obshchej redakciej Yu.P. Soboлевой, A.V. Amelichkina, N.A. Ilyuhinoj. Orel, 2023.
12. Ploennigs J., Berger M. AI art in architecture // AI in Civil Engineering. 2022. № 8.
13. Brandtzaeg P.B., Skjuve M.B., Følstad A. My AI Friend: How Users of a Social Chatbot Understand Their Human–AI Friendship // Human Communication Research. 2022.

---

**ИВЛЕВ ДМИТРИЙ ВИТАЛЬЕВИЧ** - ассистент Департамента гуманитарных наук факультета социальных наук и массовых коммуникаций, Финансовый университет при Правительстве РФ (DVIvlev@fa.ru).

**IVLEV, DMITRY V.** - Assistant of the Department of Humanities, Faculty of Social Sciences and Mass Communications, Financial University under the Government of the Russian Federation (DVIvlev@fa.ru).

---