

Социально-экономические аспекты информационного общества

**ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ КАК ОБЪЕКТ СОЦИОЛОГИЧЕСКОГО
АНАЛИЗА****Петрунин Юрий Юрьевич***Доктор философских наук, профессор**Московский государственный университет имени М.В.Ломоносова, Факультет государственного управления, заведующий кафедрой математических методов и информационных технологий в управлении**Москва, Российская Федерация**petrunin@sra.msu.ru***Нуралиева Нателма Загидовна***Московский государственный университет имени М.В.Ломоносова, Факультет государственного управления, Соискатель**Москва, Российская Федерация**natella.nuralieva@yandex.ru***Аннотация**

В настоящей статье предпринята попытка концептуализации искусственного интеллекта как объекта социологического исследования с акцентом на его современную технологическую реализацию. Выделяются две ключевые особенности технологий ИИ: агентность и техническая непрозрачность. Предлагается перспектива, в рамках которой экспертный акцент на прозрачности и объяснимости не совпадает с повседневным опытом пользователей, которые прежде всего сталкиваются с агентностью технологий. Делается вывод о необходимости дальнейшего изучения агентности ИИ и адаптации социальными науками базовых категорий – власти, справедливости и доверия.

Ключевые слова

искусственный интеллект, объяснимый ИИ, агентность, социология технологий, черный ящик

Введение

В социологии выработан самостоятельный научно-методологический инструментарий для исследования технологий. Однако появление новых технологий бросает вызов социальным наукам, заставляя переосмыслить устоявшиеся теории и подходы к исследованию. Так, распространение интернета привело к развитию цифровых методов, появлению «больших данных» и сетевых концепций общества (М. Кастельс и др.).

Сегодня искусственный интеллект (ИИ) и, в частности, генеративный ИИ, снова бросают вызов социальным наукам в силу уникальности этой технологии, ее отличиям от изучаемого ранее. И если в международной повестке исследования ИИ уже занимают заметное место в социологии технологий и смежных областях, то в российской социологии проблематика ИИ пока находится на стадии формирования [1]. Актуальность подобного анализа обусловлена стремительным распространением систем ИИ в управлении, экономике и повседневной жизни, потенциально опережающих их социологическое осмысление.

Статья носит теоретико-аналитический характер и опирается на обзор междисциплинарной литературы об ИИ, технологиях глубокого обучения и их социальном измерении. Рассматриваются предпосылки, определяющие уникальность ИИ и обозначается ключевое направление исследований, способное существенно обогатить социологическую теорию ИИ.

© Петрунин Ю. Ю., Нуралиева Н. З., 2026

Производство и хостинг журнала «Информационное общество» осуществляется Институтом развития информационного общества.

Данная статья распространяется на условиях международной лицензии Creative Commons «С указанием авторства - С сохранением условий версии 4.0 Международная» (Creative Commons Attribution – ShareAlike 4.0 International; CC BY-SA 4.0). См.

<https://creativecommons.org/licenses/by-sa/4.0/legalcode.ru>

https://doi.org/10.52605/16059921_2026_03_29

1 Технологические основания ИИ

Современная реализация идеи ИИ основана на моделях глубокого машинного обучения. В основе генеративного ИИ лежат генеративно-состязательные сети, вариационные автоэнкодеры, диффузионные модели, трансформеры и другие авторегрессионные архитектуры. [2] Видится целесообразным давать определение понятию «искусственный интеллект» через отсылку к данным технологиям, которые и делают возможной современную реализацию идеи ИИ. В этой связи имеет смысл разделять концепцию ИИ и формы ее реализации. В данной статье фокус ставится на уникальность материализованной формы ИИ. Это означает, в первую очередь, акцент на продуктах глубокого обучения и физических примерах применения этой технологии. Далее используется термин «технологии ИИ», который собирательно означает фактическую реализацию концепции ИИ с учетом его текущих технологических оснований.

2 Агентность технологий ИИ

Агентность – ключевой элемент современных социологических подходов к исследованию технологий. Под ней понимается способность к самостоятельному, намеренному, влияющему на социальную реальность действию. Б. Латур, развивая акторно-сетевую теорию (ANT), предложил идею симметрической агентности, то есть рассматривать технологии как равноправные актанты принимающие участие в конструировании социальной реальности наравне с людьми. [3]

Технологии ИИ делают эту теоретическую аналитическую конструкцию как никогда актуальной. При этом современные авторы справедливо замечают, что область агентности в социальных науках на данный момент слишком расширена и необходимо различать разные формы агентности: «Если течение воды считается «агентным» в том же смысле, что и ChatGPT, термин практически утрачивает свою информативную ценность.» [4]

Именно в системе взаимодействия актантов формируются границы автономии технологий. Если обратиться к логике устройства моделей машинного обучения, становится ясно, что создатели технологий могут намеренно или ненамеренно заложить в них некоторые смыслы, цели и идеологии. Разработчики настраивают алгоритм для удовлетворения целей бизнеса, будь то увеличение продолжительности пользователей на платформе, увеличение количества кликов, повышение стоимости покупки или сокращение затрат. Дополнительно особое значение приобретают данные: модели, обученные на эмпирических данных, воспроизводят существующие паттерны и социальные смещения (bias). [5; 6]

В результате создаются системы, влияние которых сложно переоценить. На популярной видеохостинговой платформе Youtube 70% просмотров приходится на видео, предложенные рекомендательным алгоритмом [7]. У стриминговой платформы Netflix эта цифра достигает 80% [8]. В Сбере 90% решений по сделкам оборотного кредитования для крупного и среднего бизнеса принимается с помощью ИИ [9]. Некоторые результаты работы подобных систем имеет ярко выраженную морально-этическую составляющую: система отбора персонала отклоняет заявки женщин чаще, чем мужчин [10]; транспортное средство сбивает пешехода [11]; предиктивная система обеспечения общественной безопасности чаще определяет представителя какой-либо социальной группы как потенциального правонарушителя [12]. Всё это результаты работы автономных систем.

Проблема агентности тесно связана с, и усугубляется, еще одной фундаментальной особенностью современных технологий ИИ – принципиальной непрозрачностью механизмов их работы.

3 Непрозрачность ИИ

Общество на регулярной основе взаимодействует с технологиями, внутреннюю логику которых не до конца понимает – с абстрактными системами [13]. Технологии ИИ повышают уровень абстрактности тем, что их поведение порой не может быть объяснено в том числе экспертами, самими разработчиками этих систем. Эту характеристику технологий ИИ называют «черным ящиком».

Непрозрачность технологий ИИ трактуется и как результат правовых и этических корпоративных практик, устранимая, но намеренно созданная непонятность [14], и как технически обусловленная невозможность интерпретации результатов работы системы. В отчете одного из лидеров на международном рынке технологий ИИ Anthropic, сложность интерпретации описана следующим образом: «внутреннее состояние модели – то, о чем она «думает» перед тем, как выдать свой ответ, – состоит из длинного списка чисел («активации нейронов»), которые не имеют явного значения. ... [Модель] способна понимать и оперировать широким спектром концепций, но мы не можем выявить их,

просто глядя на нейроны. Оказывается, каждая концепция представлена сразу в множестве нейронов, а каждый нейрон участвует в представлении множества концепций» [15].

Таким образом, технологии ИИ можно рассматривать как двойной черный ящик, с социальным и техническим уровнями непрозрачности.

4 Перспектива социальных наук

В нормативной и экспертной повестке вокруг ИИ проблема черного ящика стала одной из центральных тем. Преодоление этой проблемы видится как одно из необходимых условий для масштабирования и этического внедрения систем ИИ. [напр., 16] Международный регуляторный ландшафт демонстрирует возрастающий акцент именно на прозрачности и объяснимости технологии как способах обеспечения доверия [17]. Этот нарратив выходит и за пределы технических и нормативных актов: социальные науки также активно апеллируют к проблеме доверенного и объяснимого ИИ. Однако важно отметить, что идеалы прозрачности недостаточны для грамотного регулирования и настоящего понимания технологии [18], а объяснимость технологии недостаточна для доверия ей [19].

Эмпирические данные демонстрируют, что широкая общественность не осведомлена глубоко о техническом устройстве ИИ. Только 12% опрошенных по данным ВЦИОМ указали причиной недоверия, что «никто не знает как работает ИИ», и только половина опрошенных слышала и, по своей оценке, сможет объяснить, что такое технологии ИИ. [20] Этот разрыв в информированности хорошо описывается через категорию «технической неграмотности»: даже при доступности информации о работе алгоритмов, отсутствие специфических технических навыков делает системы непонятными для большинства пользователей. [21] В повседневном взаимодействии с технологией они конструируют некий её образ, отличный от экспертного понимания данной технологии [22]. На этот конструкт влияют характеристики системы другого порядка: её дизайн, фактическая корректность результатов работы систем, культурные образы, метафоры.

В данной связи напрашивается вывод, что проблема «черного ящика» оказывается скорее экспертной формулировкой проблемы, основанной на понимании технологического устройства моделей глубокого машинного обучения. В то время как общественный опыт взаимодействия с продуктами ИИ сталкивается, в первую очередь, с агентностью данной технологии – с тем, как технология «действует» по отношению к человеку в его восприятии. Именно на подобного рода взаимодействия и должны обращать особое внимание социальные науки.

Заключение

В настоящей статье предпринята попытка концептуализировать ИИ как объект исследования и обозначить перспективы для социальных наук. Выделяются две характеристики, принципиально отличающие её от иных технологий: повышенная степень социальной агентности и непрозрачность технологии. Подобная уникальность технологий ИИ требует развития аналитических подходов к изучению агентности ИИ с учётом невозможности полностью раскрыть внутреннюю логику их работы. В этих условиях особенно перспективным направлением исследований становится изучение новых форм нечеловеческой агентности, что требует критического переосмысления и адаптации фундаментальных социологических категорий — доверия, власти и справедливости — к новым социотехническим реалиям.

Литература

1. Глухих В. А., Елисеев С. М., Кирсанова Н. П. Искусственный интеллект как проблема современной социологии // Социология. Дискурс. 2022. Т. 8, № 1. С. 82–93.
2. Yazdani M., Wang S., Yang Z. et al. Generative AI in depth: A comprehensive survey on generative AI // Information Fusion. 2025. DOI: 10.1016/j.inffus.2025.102580.
3. Латур Б. Пересборка социального: введение в акторно-сетевую теорию. М.: Изд. дом Высшей школы экономики, 2014. 384 с.
4. Anicker M., Flaßhoff J., Marcinkowski F. The Matrix of AI Agency: On the Demarcation Problem in Social Theory // Sociological Theory. 2024. Vol. 42, № 4. P. 307–328.
5. O’Neil C. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York: Crown, 2016.
6. Eubanks V. Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. New York: St. Martin’s Press, 2018.

7. YouTube algorithm recommendations keep users watching // MIT Technology Review. 20.09.2022. URL: <https://www.technologyreview.com/2022/09/20/1059709/youtube-algorithm-recommendations/>
8. Plummer L. This is how Netflix's top-secret recommendation system works // Wired. 22.08.2017. URL: <https://www.wired.com/story/how-do-netflixs-algorithms-work-machine-learning-helps-to-predict-what-viewers-will-like/> (дата обращения: 26.11.2025).
9. 90% решений по сделкам оборотного кредитования для крупного и среднего бизнеса в Сбере принимается с использованием ИИ // Сбер: официальный сайт. Пресс-релиз. URL: <https://www.sberbank.com/ru/news-and-media/press-releases/article?newsID=969a88e1-f0d2-4b29-9791-3d96123164cf&blockID=7®ionID=50&lang=ru&type=NEWS> (дата обращения: 26.11.2025).
10. Dastin J. Insight – Amazon scraps secret AI recruiting tool that showed bias against women // Reuters. 10.10.2018. URL: <https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG> (дата обращения: 26.11.2025).
11. Randazzo S. Video shows moments before fatal Uber self-driving car crash // The Wall Street Journal. 21.03.2018. URL: <https://www.wsj.com/articles/video-shows-final-seconds-before-fatal-uber-self-driving-car-crash-1521673182> (дата обращения: 26.11.2025).
12. Angwin J., Larson J., Mattu S., Kirchner L. Machine bias: There's software used across the country to predict future criminals. And it's biased against Blacks // ProPublica. 23.05.2016. URL: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing> (дата обращения: 26.11.2025).
13. Гидденс Э. Последствия современности. Пер. с англ.; вступ. ст. Т. А. Дмитриева. – М.: Праксис, 2011. – 352 с.
14. Pasquale F. The Black Box Society: The Secret Algorithms That Control Money and Information. Cambridge, MA: Harvard University Press, 2015. 320 p.
15. Anthropic. Mapping the mind of a large language model. 2024. URL: <https://www.anthropic.com/research/mapping-mind-language-model> (дата обращения: 26.11.2025).
16. Mathew D.E., Ehem D.U., Ikegwu A.C., Ukeoma P.E., Dibiazue N.F. Recent Emerging Techniques in Explainable Artificial Intelligence to Enhance the Interpretable and Understanding of AI Models for Human // Neural Processing Letters. 2025. Vol. 57, № 1. Art. 16. DOI: 10.1007/s11063-025-11732-2.
17. Lund B.D., Orhan Z., Mannuru N.R. et al. Standards, Frameworks, and Legislation for Artificial Intelligence (AI) Transparency // AI and Ethics. 2025. Vol. 5. P. 3639–3655. DOI: 10.1007/s43681-025-00661-4.
18. Ananny M., Crawford K. Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability // New Media & Society. 2018. Vol. 20, № 3. P. 973–989.
19. Dahlin E. Trust in AI // AI & Society. 2025. Vol. 40. P. 6089–6095. DOI: 10.1007/s00146-025-02429-0.
20. Доверие к ИИ // ВЦИОМ. Аналитический обзор. 2024. URL: <https://wciom.ru/analytical-reviews/analiticheskii-obzor/doverie-k-ii> (дата обращения: 26.11.2025)
21. Burrell J. How the machine 'thinks': Understanding opacity in machine learning algorithms // Big Data & Society. 2016. Vol. 3, № 1. P. 1–12.
22. Fortuna P., Gorbaniuk O. What Is Behind the Buzzword for Experts and Laymen: Representation of «Artificial Intelligence» in the IT-Professionals' and Non-Professionals' Minds // Europe's Journal of Psychology. 2022. Vol. 18, № 2. P. 207–218. DOI: 10.5964/ejop.5473.

ARTIFICIAL INTELLIGENCE AS AN OBJECT OF SOCIOLOGICAL INQUIRY

Petrinin, Yuriy Yurievich

DSc in philosophy, professor

Lomonosov Moscow State University, School of public administration, Department of mathematical methods and information technology in management, chairman

Moscow, Russia

petrulin@spa.msu.ru

Nuralieva, Natella Zagidovna

Lomonosov Moscow State University, School of public administration, PhD candidate

Moscow, Russia

natella.nuralieva@yandex.ru

Abstract

This article seeks to conceptualize artificial intelligence as an object of sociological inquiry, with a focus on its contemporary technological manifestations. It highlights two key features of AI technologies: agency and technical opacity. The article proposes an analytical perspective in which the expert emphasis on transparency and XAI is contrasted with the everyday experience of users who primarily encounter the agency of these technologies. It concludes by arguing for the need to further study AI agency and to adapt core categories of social sciences – power, justice, and trust – to new socio-technical configurations of AI.

Keywords

artificial intelligence, explainable AI, agency, sociology of technology, black box

References

1. Glukhikh, V. A., Eliseev, S. M., Kirsanova, N. P. *Iskusstvennyi intellekt kak problema sovremennoi sotsiologii // Sotsiologiya. Diskurs. 2022. T. 8, № 1. S. 82–93.*
2. Yazdani M., Wang S., Yang Z. et al. *Generative AI in depth: A comprehensive survey on generative AI // Information Fusion. 2025. DOI: 10.1016/j.inffus.2025.102580.*
3. Latour, B. *Peresborka sotsial'nogo: vvedenie v aktorno-setevuyu teoriyu (Reassembling the Social: An Introduction to Actor-Network-Theory). M.: Izd. dom Vysshei shkoly ekonomiki, 2014. 384 s.*
4. Anicker M., Flaßhoff J., Marcinkowski F. *The Matrix of AI Agency: On the Demarcation Problem in Social Theory // Sociological Theory. 2024. Vol. 42, № 4. P. 307–328.*
5. O'Neil C. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. New York: Crown, 2016.*
6. Eubanks V. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor. New York: St. Martin's Press, 2018.*
7. *YouTube algorithm recommendations keep users watching // MIT Technology Review. 20.09.2022. URL: <https://www.technologyreview.com/2022/09/20/1059709/youtube-algorithm-recommendations/>*
8. *Plummer L. This is how Netflix's top-secret recommendation system works // Wired. 22.08.2017. URL: <https://www.wired.com/story/how-do-netflixs-algorithms-work-machine-learning-helps-to-predict-what-viewers-will-like/>*
9. *90% reshenii po sdelkam oborotnogo kreditovaniya dlya krupnogo i srednego biznesa v Sbere prinimaetsya s ispol'zovaniem II // Sber: ofitsial'nyi sait. Press-reliz. URL: <https://www.sberbank.com/ru/news-and-media/press-releases/article?newsID=969a88e1-f0d2-4b29-9791-3d96123164cf&blockID=7®ionID=50&lang=ru&type=NEWS>*
10. *Dastin J. Insight – Amazon scraps secret AI recruiting tool that showed bias against women // Reuters. 10.10.2018. URL: <https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG>.*
11. *Randazzo S. Video shows moments before fatal Uber self-driving car crash // The Wall Street Journal. 21.03.2018. URL: <https://www.wsj.com/articles/video-shows-final-seconds-before-fatal-uber-self-driving-car-crash-1521673182>.*

12. Angwin J., Larson J., Mattu S., Kirchner L. Machine bias: There's software used across the country to predict future criminals. And it's biased against Blacks // ProPublica. 23.05.2016. URL: <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
13. Giddens, A. *Posledstviya sovremennosti (The Consequences of Modernity)*. Per. s angl.; vstup. st. T. A. Dmitrieva. M.: Praxis, 2011. 352 s.
14. Pasquale F. *The Black Box Society: The Secret Algorithms That Control Money and Information*. – Cambridge, MA: Harvard University Press, 2015. 320 p.
15. Anthropic. Mapping the mind of a large language model. 2024. URL: <https://www.anthropic.com/research/mapping-mind-language-model>.
16. Mathew D.E., Ebem D.U., Ikegwu A.C., Ukeoma P.E., Dibiaezue N.F. Recent Emerging Techniques in Explainable Artificial Intelligence to Enhance the Interpretable and Understanding of AI Models for Human // *Neural Processing Letters*. 2025. Vol. 57, № 1. Art. 16. DOI: 10.1007/s11063-025-11732-2.
17. Lund B.D., Orhan Z., Mannuru N.R. et al. Standards, Frameworks, and Legislation for Artificial Intelligence (AI) Transparency // *AI and Ethics*. 2025. Vol. 5. P. 3639–3655. DOI: 10.1007/s43681-025-00661-4.
18. Ananny M., Crawford K. Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability // *New Media & Society*. 2018. Vol. 20, № 3. P. 973–989.
19. Dahlin E. Trust in AI // *AI & Society*. 2025. Vol. 40. P. 6089–6095. DOI: 10.1007/s00146-025-02429-0.
20. Doverie k II // *Vserossiiskii tsentr izucheniya obshchestvennogo mneniya (VTsIOM)*. Analiticheskii obzor. 2024. URL: <https://wciom.ru/analytical-reviews/analiticheskii-obzor/doverie-k-ii>
21. Burrell J. How the machine 'thinks': Understanding opacity in machine learning algorithms // *Big Data & Society*. 2016. Vol. 3, № 1. P. 1–12.
22. Fortuna P., Gorbaniuk O. What Is Behind the Buzzword for Experts and Laymen: Representation of «Artificial Intelligence» in the IT-Professionals' and Non-Professionals' Minds // *Europe's Journal of Psychology*. 2022. Vol. 18, № 2. P. 207–218. DOI: 10.5964/ejop.5473.