

Секция «9.14 Большие данные и искусственный интеллект в современном
государственном управлении»

**Дискурс конкуренции человека и искусственного интеллекта в формировании
доверия при внедрении ИИ в государственное управление**

Научный руководитель – Петрунин Юрий Юрьевич

Нуралиева Нателла Загидовна

Аспирант

Московский государственный университет имени М.В.Ломоносова, Факультет
государственного управления, Москва, Россия

E-mail: natella.nuralieva@yandex.ru

Технологии искусственного интеллекта (далее, ИИ) проникают в системы государственного управления: они централизованно внедряются ведомствами и могут выходить на передовую во взаимодействиях с гражданами. При этом одним из ключевых факторов эффективного внедрения и принятия технологии является общественное доверие к ней. В свою очередь, доверие формируется под влиянием общественного дискурса и существующих представлений о технологии.

Одним из ключевых мотивов современной новостной и академической повестки вокруг ИИ выступает образ конкуренции между технологией и человеком. Это транслируется широким массам, например, в новостях о массовых сокращениях и росте требований работодателей к цифровым навыкам и навыкам работы с ИИ. Подобный дискурс формирует логику общественного обсуждения ИИ: технология начинает восприниматься не просто как инструмент, а как потенциальный субъект принятия решений, которому можно предпочесть человека.

Академический мир тоже изучает проявляемые обществом предпочтения между технологией и человеком [4, 5] и связи недоверия человеку с доверием технологиям [напр., 2].

Суждения о конкуренции технологии и человека могут быть правомерны и обоснованы в связи с повышенной агентностью и антропоморфностью технологий ИИ, принятием ими участия в принятии решений, будучи встроенными в соответствующие процессы. Вместе с тем, подобного рода дискурс влечет и смещение критериев, по которым технология оценивается как заслуживающая доверия, своего рода трансформацию доверия к технологии.

Одним из механизмов, делающих образ конкуренции возможным, выступает перенос привычных для межличностного взаимодействия образов и интерпретаций на ИИ. Еще задолго до появления столь антропоморфных технологий как большие языковые модели исследования показывали, что люди склонны применять к автоматизированным системам и алгоритмам социальные ожидания и модели доверия [6, 7]. Это означает перенос конкретных ожиданий: компетентности, доброжелательности, способности принимать обоснованные решения. Ожидания, ранее свойственные межличностным отношениям, начинают адресоваться и технологиям, дополняя более привычные для них ожидания надежности и бесперебойности работы и тем самым углубляя доверие к ним. Современные авторы идут дальше и утверждают, что дальнейшее усиление антропоморфизации технологий может повлечь уменьшение роли создателей технологии при формировании мнения обществом об ИИ и усиление его восприятия как социального актора [3].

Однако важно понимать, что перенос межличностной логики доверия возможен, скорее, функционально. Например, речь может идти об ИИ в ситуациях поддержки принятия решений (DSS) [6]. Применение межличностной логики доверия к профессиональным ролям, например к представителям государственных учреждений, имеет менее прямолинейный характер. Человек в профессиональном контексте, будучи представителем профессии

или института, воспринимается как элемент более широкой системы, и его оценка неизбежно связана с ассоциациями, свойственными этой системе.

Более того, сама технология – искусственный интеллект – представляет собой сложную социотехническую систему, доверие к которой формируется на основе множества факторов и широкого набора соприкасающихся с технологией адресатов доверия: разработчиков, инвесторов, регуляторов, лидеров мнений, и других вовлеченных акторов [1].

В этой связи важно понимать, что межличностное доверие – недостаточная модель для объяснения доверия к самим технологиям ИИ. Технологии и эксперты не являются обособленными конкурирующими сущностями. Они встроены в социотехнические и профессиональные системы, оперирующие в соответствии с определенными логикой, нормами, процедурами. А значит, сравнение ИИ с человеком не может быть достаточным: оно хорошо объясняет часть механизма доверия, но не всю его социальную и институциональную природу. Механизм межличностного доверия – лишь один из уровней доверия к технологии, который безусловно требует изучения, однако с учетом контекста его взаимодействия с другими уровнями доверия. Именно взаимосвязь уровней доверия должна изучаться для формирования комплексной модели доверия к ИИ как сложной социотехнической системе.

Таким образом, при внедрении ИИ в государственное управление необходимо учитывать как склонность общества осмысливать ИИ через логику конкуренции с человеком, так и ограниченность такого подхода для понимания доверия к сложным социотехническим системам.

Источники и литература

- 1) Гарбук С.В. Модель доверия к прикладным системам искусственного интеллекта // Вестник Московского университета. Серия 21. Управление (государство и общество). 2024. Т. 21. №. 4. С. 151–169. DOI: 10.55959/MSU2073-2643-21-2024-4-151-169.
- 2) Gerlich M. Exploring Motivators for Trust in the Dichotomy of Human–AI Trust Dynamics // Social Sciences. 2024. Vol. 13. No. 5. Article 251. DOI: 10.3390/socsci13050251.
- 3) Lalot F., Bertram A.-M. When the Bot Walks the Talk: Investigating the Foundations of Trust in an Artificial Intelligence (AI) Chatbot // Journal of Experimental Psychology: General. 2025. Vol. 154. No. 2. P.533–551. DOI: 10.1037/xge0001696.
- 4) Langer M., König C.J., Back C., Hemsing V. Trust in Artificial Intelligence: Comparing Trust Processes Between Human and Automated Trustees in Light of Unfair Bias // Journal of Business and Psychology. 2023. Vol. 38. No. 3. P. 493–508. DOI: 10.1007/s10869-022-09829-9.
- 5) Logg J.M., Minson J.A., Moore D.A. Algorithm Appreciation: People Prefer Algorithmic to Human Judgment // Organizational Behavior and Human Decision Processes. 2019. Vol. 151. P. 90–103. DOI: 10.1016/j.obhdp.2018.12.005.
- 6) Madhavan P., Wiegmann D.A. Similarities and Differences Between Human–Human and Human–Automation Trust: An Integrative Review // Theoretical Issues in Ergonomics Science. 2007. Vol. 8. No. 4. P. 277–301. DOI: 10.1080/14639220500337708.
- 7) Nass C., Steuer J., Tauber E.R. Computers Are Social Actors // Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. N.Y., 1994. P. 72–78. DOI: 10.1145/191666.191703.