

Секция «Искусственный интеллект и цифровая трансформация в бизнесе и государственном управлении»

Человеческая этика и искусственный интеллект: на стыке технологий и морали

Научный руководитель – Бахриев Бахри Хуршедович

Романова Кристина Витальевна

Студент (бакалавр)

Московский городской университет управления Правительства Москвы, Москва, Россия

E-mail: k_romanova_2003@mail.ru

Четвертая промышленная революция, связанная с цифровыми технологиями, коренным образом меняет новый общественный уклад[3]. Цифровая трансформация затрагивает государственные учреждения, предприятия и повседневные аспекты жизни каждого из нас. Прорывными стали технологии искусственного интеллекта (ИИ) ответ на развитие которых возвращает человечество к этическим принципам, к переосмыслению представления о том, что является моральным или аморальным в различных ситуациях[7]. В этом контексте изучение этических аспектов применения ИИ становится актуальной научной и практической задачей.

Изучение взаимоотношения этики и ИИ требует категориального уточнения. Термин «этика» было предложено Аристотелем, под которым он понимал науку о правильном поведении. Термин «искусственный интеллект» впервые в 1965 году предложил Джон Маккарти. В этом отношении существует ключевое различие между слабым (или узким) интеллектом и сильным. Слабый искусственный интеллект относится к системам, предназначенным для выполнения конкретных задач, таких как технологии «умного дома» или инструменты онлайн-перевода, но не способным выполнять более широкие когнитивные функции. И наоборот, для сильного ИИ характерны способность к самостоятельному обучению, решению различных интеллектуальных задач и самосознанию.

Этика искусственного интеллекта, как считают исследователи, состоит из двух аспектов[2]. Первый аспект показывает, что изначально программа ИИ была написана людьми, но позже система ИИ работала практически независимо и могла совершенствовать себя и улучшать свои параметры. Трудность заключается в том, что моральный выбор определяется не четко определенными юридическими нормами, а сформулированными правилами, принципами, личными мнениями и всеми субъективными решениями, которые могут быть оценены как «хорошие» или «плохие». Второй аспект этики искусственного интеллекта включает анализ и предотвращение этических конфликтов, возникающих в процессе использования искусственного интеллекта. К ним относятся: вторжение в частную жизнь, возможная дискриминация, социальное расслоение и даже проблемы с трудоустройством.

Учитывая появление технологий, которые могут быть использованы в заранее обдуманных целях, этические аспекты искусственного интеллекта стали особенно актуальными. Например, технология *deepface* позволяет манипулировать видеоконтентом, заменяя изображение одного человека изображением другого, что вызывает серьезные опасения по поводу кражи личных данных, дезинформации и личной безопасности. С популярностью систем искусственного интеллекта возникают новые социальные риски и проблемы, связанные с человеческой идентичностью и неприкосновенностью частной жизни. Прогнозирование и предотвращение этих рисков является важной частью ответственного технологического развития[6].

Для решения этих задач в России был принят кодекс этики в области искусственного интеллекта в 2021 году.[1] В кодексе подчеркивается важность ответственности человека

за этические риски, связанные с технологией искусственного интеллекта. Он выступает за развитие искусственного интеллекта в направлении, ориентированном на людей, чтобы гарантировать, что эти разработки улучшат, а не помешают человеческим способностям, когнитивным функциям и защите культуры. Эти моральные принципы направлены на установление баланса между техническим прогрессом и основными правами человека.

Некоторые государства приняли документы, непосредственно связанные с этикой искусственного интеллекта в стране (Австралия[4]). В США в 2021 году Счетная палата вынесла рекомендации по системе подотчетности государственных учреждений в области использования искусственного интеллекта, включая рекомендации по применению этических принципов искусственного интеллекта в определенных областях.[5]

В конечном счете, пересечение человеческой этики и искусственного интеллекта побуждает человечество задуматься о его социальных ценностях. Живя в быстроменяющемся мире, человек должен быть бдительным в отношении моральных последствий своего технологического выбора. Только такое обеспечение соответствия инноваций этическим принципам и способствовать гармоничному сосуществованию человечества и создаваемых им интеллектуальных систем.

Источники и литература

- 1) Сазонова М. Искусственный интеллект и этика: эксперты оценили проект нового кодекса / М. Сазонова // Гарант.ру. – 2021. <http://www.garant.ru/article/1487350/>
- 2) Шляпников В.В. Некоторые проблемы этики искусственного интеллекта // Идеи и идеалы. 2023. №15 (2-2). С. 365-376
- 3) Шумский С. Воспитание машин: новая история разума / С. Шумский ; под редакцией С. Филоновича, П. Суворовой. — Москва : Альпина нон-фикшн, 2024. — 174 с.
- 4) Australia's AI Ethics Principles // Режим доступа: URL: <https://www.industry.gov.au/data-and-publications/australias-artificial-intelligence-ethics-framework/australias-ai-ethics-principles>
- 5) Artificial Intelligence. An Accountability Framework for Federal Agencies and Other Entities. GAO 2021. <https://www.gao.gov/assets/gao-21-519sp.pdf>
- 6) Hauer T. Importance and limitations of AI ethics in contemporary society // Humanities and Social Sciences Communications. 2022. № 9 (272). DOI: <https://doi.org/10.1057/s41599-022-01300-7>
- 7) Munn L. The uselessness of AI ethics // AI Ethics. 2023. Vol. 3. P. 869–877