

Этическое регулирование технологий искусственного интеллекта

Н. Ю. Ключева¹

В сообщении речь пойдет об этико-прикладных проблемах создания и использования технологий искусственного интеллекта. Рассматриваются существующие этические кодексы, регламентирующие сферу искусственного интеллекта и их ключевые принципы.

Ключевые слова: искусственный интеллект, философия техники, прикладная этика

Искусственный интеллект, понимаемый как область компьютерных наук, которая занимается созданием программ и компьютеров, имитирующих когнитивные функции человека и способных проявлять разумное поведение, имеет на настоящий момент результаты как в теоретической сфере, так и в области разработки и применения конкретных инновационных технологий.

Причины, по которым задача этико-правового регулирования процессов разработки и использования технологий искусственного интеллекта на национальном и международном уровне является приоритетной, включает принципиальную новизну некоторых аспектов этих технологий: автономность систем в области принятия решений; непрозрачность принципов принятия решений. Отдельный аспект этических вопросов связан с такой особенностью некоторых систем искусственного интеллекта как обучение на массиве данных. Тестирование систем на основе технологий искусственного интеллекта и реальная практика их применения показывают, что обучение на массиве данных может привести к ошибкам и дискриминации, в основе которых лежит несовершенство самого массива. Целый ряд вопросов этического характера связан с прогностической направленностью задач, которые призваны решать системы искусственного интеллекта. Довольно остро этико-прикладные вопросы применения технологий искусственного интеллекта проявляются в таких социально значимых сферах общественной жизни, как медицина и здравоохранение, государственное управление и образование. В этих областях этические вопросы традиционно играют большую роль, при этом внимание общественности к соблюдению принципов справед-

¹ Ключева Наталья Юрьевна — доцент философского ф-та МГУ, e-mail: Klyueva.msu@gmail.com.

Klyueva Natalia Yurievna — associate professor, Faculty of Philosophy, Lomonosov Moscow State University

ливости, ответственности за принятые решения, и прозрачности работы максимально высоко.

Попытка представить базовый этический стандарт регулирования технологий искусственного интеллекта сделана в «Первом проекте рекомендаций об этических аспектах искусственного интеллекта» ЮНЕСКО [1]. Документ призван зафиксировать базовые ценностные установки искусственной интеллектуальной системы, среди которых уважение, защита и поощрение человеческого достоинства, прав человека и основных свобод; благополучие окружающей среды и экосистем; обеспечение разнообразия и инклюзивности; поощрение мира и согласия. В РФ также разработаны этические рекомендации по использованию ИИ. Самым масштабным на сегодняшний день является «Кодекс этики в сфере искусственного интеллекта», который был подписан в октябре 2021 года [2]. Человеко-ориентированный подход в создании и реализации ИИ, зафиксированный в документе, основывается на схожих ценностных ориентирах, что и проект ЮНЕСКО. В числе ключевых задач: контроль за соответствием закону, за соблюдением принципа уважения свободы воли человека, за отсутствием дискриминации, за оценкой рисков, безопасностью данных и информационной безопасностью, включая принцип подконтрольности ИИ человеку и отказ от передачи ИИ полномочий ответственного нравственного выбора. Несмотря на отказ ИИ в моральной субъектности, рост использования технологии ИИ будет приводить к трансформации понимания принципов распределения ответственности между машиной и человеком. Соблюдение указанных в документе принципов предполагается в добровольном порядке и связывается участниками с категорией социальной ответственности технологических компаний, компаний-разработчиков. Представляется, что для полноценного этического регулирования развивающейся сферы ИИ требуется включение документов этического регулирования в национальную программу развития ИИ на государственном уровне, что позволит минимизировать влияние коммерческих интересов в данном процессе.

Список литературы

- [1] *Report «First draft of the Recommendation on the Ethics of Artificial Intelligence»*, 2020, <https://unesdoc.unesco.org/ark:/48223/pf0000373434>.
- [2] *Кодекс этики в сфере искусственного интеллекта*, 2021, <https://www.aiethic.ru/code?fbclid=IwAR2uavuwQ8icTAppyER42IjpvnfHzMo8kuEnjx5OzSZItPFPV3Bp8oJFNQ>.

Ethical regulation of artificial intelligence technologies **Klyueva N.Y.**

The message will focus on ethical and applied problems of creating and using artificial intelligence technologies. Existing ethical codes regulating the field of artificial intelligence and their key principles are considered.

Keywords: artificial intelligence, philosophy of technology, applied ethics

References

- [1] *Report «First draft of the Recommendation on the Ethics of Artificial Intelligence»*, 2020, <https://unesdoc.unesco.org/ark:/48223/pf0000373434>.
- [2] *Code of Ethics for Artificial Intelligence*, 2021 (Russian), <https://www.aiethic.ru/code?fbclid=IwAR2ua-vuwQ8icTApgyER42IjpvnfHzMo8kuEnjx5OzSZItPFPV3Bp8oJFNQ>.