

УДК 378.1

Афанасьева Инга Геннадьевна

старший преподаватель

Чермашенцева Мария Андреевна

студентка

Пикулина Екатерина Александровна

студентка

ФГБОУ ВО «Томский государственный университет
систем управления и радиоэлектроники»

г. Томск, Томская область

DOI 10.21661/r-562240

ЭТИЧЕСКИЕ ПРИНЦИПЫ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА КАК ОСНОВА ПОДГОТОВКИ СПЕЦИАЛИСТОВ СФЕРЫ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ

Аннотация: в настоящее время общество сталкивается с большим объемом информации, который непрерывно генерируется в различных сферах деятельности общества. Эффективным инструментом по работе с информацией специалистами в области ИТ определяется технология искусственного интеллекта. В то же время общество имеет негативный опыт с позиции этики применения данной технологии. Это определило необходимость модернизации образовательных программ будущих специалистов сферы ИТ, учитывающую этические принципы в области разработки и применения искусственного интеллекта.

Ключевые слова: искусственный интеллект, этические принципы искусственного интеллекта, подготовка специалистов, социальная ответственность, информационные технологии.

Современный мир характеризуется постоянной генерацией большого количества информации во всех сферах деятельности общества. Информация стала ключевым ресурсом функционирования как социальных, так и экономических систем, поэтому общество в постоянном поиске наиболее эффективного

инструмента обработки и анализа информации [1]. Актуальной технологией для решения данного вопроса может рассматриваться искусственный интеллект (ИИ). Преимуществом данной технологии заключается в том, что она позволяет создавать информационные системы, адаптирующиеся к внешним условиям. ИИ помогает автоматизировать рутинную работу, анализировать большие массивы данных, прогнозировать изменения спроса и предложения, оптимизировать процессы производства и управления запасами [4].

Однако, несмотря на свои многочисленные преимущества, алгоритмы искусственного интеллекта могут содержать ошибки и неучтенные критерии, которые значительно влияют на жизнь общества. Например, в 2016 году произошло первое дорожно-транспортное происшествие с летальным исходом, виновником которого был автопилот автомобиля Tesla Model S. Автопилот не смог различить белый грузовик с прицепом и попытался проехать под прицепом, тем самым спровоцировав аварию [8]. Также в 2016 году чат-бот со встроенным ИИ, созданный с целью развлечения, научился расизму, национализму и сексизму от людей несмотря на встроенный фильтр вульгарных выражений [7]. В 2021 году организация объединенных наций впервые зафиксировало убийство людей с использованием военных дронов с встроенным ИИ без прямого приказа оператора [2].

В момент, когда возникают последствия от недостатков и ошибок работы алгоритмов искусственного интеллекта возникает вопрос: кто несет ответственность за результат работы и применение искусственного интеллекта? Самообучающаяся информационная система или человек, создавший алгоритм, на основе которого данная система принимает решение, и какие этические принципы следует использовать, чтобы предотвратить подобные случаи в будущем?

Вопросы этики ИИ начали подниматься еще в 19 веке. В 1950 году Алан Тьюринг уже анализировал последствия создания «думающих машин». В частности, он рассматривал вопрос о том, что машины не могут совершать ошибки. Также отмечал, что ошибки могут возникать из-за неправильных или некорректных исходных данных, и тем не менее, машины выполняют все математические операции точно [5].

В современности, разработчики искусственного интеллекта должны стремиться к созданию систем, которые максимально снижают вероятность ошибок и обеспечивают безопасность для пользователей. Должна быть сформирована ответственность за детальное тестирование искусственного интеллекта, а также за обновление и усовершенствование алгоритмов и систем с учетом обратной связи, опыта и рисков.

Поскольку ИИ может принимать решения, влияющие на безопасность жизни людей, возникают вопросы о качестве и границах действия алгоритмов искусственного интеллекта, а также о распределении ответственности за принятые решения. Этот вопрос является актуальным и для РФ, так как развитие проекта «Искусственный интеллект», утвержденный в 2021 году для реализации Национальной стратегии РФ, подразумевает применение ИИ в 95% всех компаний Российской Федерации к 2030 году [3]. В связи с этим, государство, должно принять активное участие в разработке этических принципов искусственного интеллекта, создать нормативные акты и правила, которые определяют ожидания и требования к использованию искусственного интеллекта, чтобы снизить риски ненадлежащего использования инструментов ИИ. Цель таких нормативных актов и правил состоит в том, чтобы определить рамки и правила использования искусственного интеллекта. Эти принципы могут включать в себя вопросы конфиденциальности данных, прозрачности работы алгоритмов, ответственности за принимаемые решения и другие аспекты.

В России 26 октября в 2021 году был разработан документ, регламентирующий применения технологии искусственного интеллекта «Кодекс этики в сфере искусственного интеллекта» [6]. Документ также носит рекомендательный характер и содержит общие принципы этики и стандарты поведения, но отражает вопросы о необходимости определения зон ответственности при разработке и использовании ИИ, осознание рисков использования ИИ, рамок использования ИИ. В кодексе описана рекомендация поддерживать должный уровень профессиональных компетенций разработчиков информационных систем, необходимый для безопасной и эффективной работы с ИИ, содействовать повышению профессиональной компетенции специалистов в области ИИ, в том числе в составе образовательных программ

высших учебных заведений. Для нивелирования рисков, связанных с внедрением искусственного интеллекта во все сферы деятельности общества, образовательные программы подготовки специалистов сферы информационных технологий должны включать вопросы этики ИИ, к которым относятся:

- главный приоритет развития технологий ИИ в защите интересов и прав людей и отдельного человека;
- осознание ответственности при создании и использовании ИИ;
- ответственность за последствия применения ИИ всегда несет человек;
- технологии ИИ нужно внедрять там, где это принесёт пользу людям;
- интересы развития технологий ИИ выше интересов конкуренции;
- максимальная прозрачность и правдивость в информировании об уровне развития технологий ИИ, их возможностях и рисках [6].

Специалисты сферы информационных технологий должны четко понимать важность и значимость данной технологии для создания систем, обеспечивающих безопасность, защиту данных и личной информации, также должны взаимодействовать между собой, совместно повышать компетенции, пользоваться и разрабатывать корректные методы сравнения и разработки данных систем. Кроме того, к задачам данных специалистов необходимо отнести просвещение пользователей о развитии искусственного интеллекта, и сопутствующих рисках его применения в профессиональной и повседневной жизни.

Таким образом, рассматривая технологию искусственного интеллекта как инновационный инструмент, важно ответственно и осознанно подходить к работе с ним. Вопросы этики искусственного интеллекта должны рассматриваться в процессе профессиональной подготовки специалистов в сфере информационных технологий, обеспечивая тем самым безопасное внедрение, анализ рисков применения технологии искусственного интеллекта при создании и внедрении информационных систем.

Список литературы

1. Герасимов В.А. Проблемы обработки больших данных [Электронный ресурс] / В.А. Герасимов, С.В. Шайтура, А.Д. Емельянов [Электронный ресурс]. –

Режим доступа: <https://ies.unitech-mo.ru/files/upload/publications/15844/07dc80e397ed6b0e6dba68822cd3b831.pdf>

2. Впервые в истории боевой робот убил человека по собственной инициативе [Электронный ресурс]. – Режим доступа: https://www.cnews.ru/news/top/2021-05-31_vpervye_v_istorii_boevoj

3. Выступление вице-премьера Дмитрия Чернышенко по итогам стратегической сессии по развитию ИИ до 2030 года [Электронный ресурс]. – Режим доступа: <https://tass.ru/ekonomika/18846061>

4. Калашникова И.В. Использование искусственного интеллекта при принятии управленческих решений в проектном менеджменте / И.В. Калашникова, Д.В. Несмеянов [Электронный ресурс] – Режим доступа: <https://cyberleninka.ru/article/n/ispolzovanie-iskusstvennogo-intellekta-pri-prinyatii-upravlencheskih-resheniy-v-proektnom-menedzhmente/viewer>

5. Карпов В.Э. К вопросу об этике и системах искусственного интеллекта / В.Э. Карпов, П.М. Готовцев, Г.В. Ройзензон // Философия и общество. – 2018. – 105 с.

6. Кодекс этики в сфере искусственного интеллекта [Электронный ресурс]. – Режим доступа: https://ethics.a-ai.ru/assets/ethics_files/2023/04/05/Кодекс_этики_в_сфере_ИИ_финальный_1.pdf

7. Овечкин О.И. Чатбот от Microsoft научился ненавидеть евреев и соглашаться с Гитлером / О.И. Овечкин [Электронный ресурс]. – Режим доступа: <https://rb.ru/news/poor-tay/>

8. Tesla driver dies in first fatal crash while using autopilot mode [Электронный ресурс]. – Режим доступа: <https://www.theguardian.com/technology/2016/jun/30/tesla-autopilot-death-self-driving-car-elon-musk>