

## Глобальные инициативы по регулированию Искусственного интеллекта

Научный руководитель – Шестова Татьяна Львовна

*Васильева Елена Игоревна*

*Студент (магистр)*

Московский государственный университет имени М.В.Ломоносова, Геологический факультет, Кафедра геохимии, Москва, Россия

*E-mail: cntqpb@mail.ru*

Искусственный интеллект – новая реальность общества будущего [3]. Используя технологические возможности «мыслящих машин» и передавая им часть своих социальных функций, человек, с одной стороны, многократно усиливает свои внутренние ресурсы, с другой – оказывается в потенциальной зависимости от перспектив развития искусственного интеллекта.

Теория искусственного интеллекта уже более полутора десятилетий обращается к этическим проблемам, связанным с машинным сознанием. Одна из таких проблем – необходимость принятия этических кодексов, ограничивающих деятельность разработчиков систем искусственного интеллекта. Применение подобных кодексов должно иметь международный масштаб, так как влияние новых технологий на все стороны жизни общества и человека не может быть ограничено национальными рамками. Искусственный интеллект – общее детище человечества, и «воспитывать», «социализировать» его надо сообща.

На международном уровне принято большое количество документов, регламентирующих развитие и функционирование искусственного интеллекта, однако большинство из них носит только рекомендательный характер. Такими документами являются, например, Ethics guidelines for trustworthy AI [4], разработанный в ЕС в 2019 году, или материалы Глобального партнерства по искусственному интеллекту, действующего с 2020 года под эгидой Организации экономического сотрудничества и развития.

Еще одной структурой, обратившейся к этическим нормам, регулирующим ИИ, стала ЮНЕСКО. В 2021 году ЮНЕСКО приняла документ, подписанный 194 странами-участницами, Recommendation on the Ethics of Artificial Intelligence, в котором обозначила основные ценности и принципы цифровой политики, направленной на устойчивое будущее нашей цивилизации [5].

В ООН риски, связанные с развитием искусственного интеллекта, на постоянной основе обсуждаются с 2018 года, когда была создана Группа высокого уровня по цифровому сотрудничеству Генерального секретаря ООН. Группа занимается глобальными аспектами цифрового развития как регионов и стран, так и мира в целом.

В мае 2020 года в ООН была представлена «Дорожная карта по цифровому сотрудничеству: осуществление рекомендаций Группы высокого уровня», доклад по основному содержанию которой был сделан Генеральным секретарем А.Гутерришем. В докладе рассматривается текущее состояние и перспективы развития цифровых технологий на планете, указываются проблемы, связанные с цифровым неравенством, новыми рисками, безопасностью и др. Рассматриваются и проблемы искусственного интеллекта. В докладе говорится: «Искусственный интеллект может также существенно подорвать безопасность пользователей по всему миру и их способность самостоятельно воздействовать на окружающую действительность» [1]. Особо упоминаются угрозы, связанные с автономными системами летального действия, способными принимать решения о жизни и смерти, без вмешательства человека, и сети, производящие «дипфейки», направленные на манипулирование общественным сознанием. [1 ]

1 января 2025 года Генеральной Ассамблеей ООН было создано Управление Организации Объединенных Наций по цифровым и новым технологиям (ODET) [2], в задачу которого входит координация мировых усилий по переходу на цифровые технологии в большинстве социально значимых областей и в том числе – международное регулирование развития искусственного интеллекта.

Таким образом, человечество в целом осознает риски и угрозы, связанные с развитием искусственного интеллекта, однако сотрудничество в совместном купировании этих угроз пока остается на очень слабом уровне. Принятые документы носят лишь рекомендательный характер, хотя наметившаяся координация усилий под эгидой авторитетных международных организаций позволяет надеяться на осуществление продуктивного диалога в глобальном масштабе и в конечном счете – на решение вопроса в интересах человека и человечества.

### Источники и литература

- 1) 1. Дорожная карта Генерального секретаря по цифровому сотрудничеству // <http://docs.un.org/ru/A/74/821> (дата обращения: 25.02.2026).
- 2) 2. Развитие сотрудничества в области управления технологиями ради лучшего будущего // <https://www.un.org/digital-emerging-technologies/ru> (дата обрац. 20.02.2026)
- 3) 3. Садовничий В. А., Акаев А. А., Ильин И. В., Малков С. Ю., Коротяев А. В., Гринин Л. Е., Андреев А. И. Общество будущего и переход к нему // История и современность. 2023, № 1 (47)
- 4) 4. Ethics guidelines for trustworthy AI [Электронный ресурс] URL: <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai> (дата обращения: 25.02.2026).
- 5) 5. Recommendation on the Ethics of Artificial Intelligence [Электронный ресурс] URL: <https://unesdoc.unesco.org/ark:/48223/pf0000381137> (дата обращения: 25.02.2026).