



УНИВЕРСИАДА ПО ИННОВАТИКЕ. Отборочный этап (заочный).

КЕЙС «ЭТИКА ПРИМЕНЕНИЯ ИИ: МЕЖДУ ПРОГРЕССОМ И ИНФОРМАЦИОННОЙ БЕЗОПАСНОСТЬЮ»

Сегодня искусственный интеллект (ИИ) становится технологией, которая меняет саму структуру общества и экономики. По данным McKinsey Global Institute, внедрение ИИ способно увеличить мировой ВВП на 15,7 трлн долл к 2030 году, а Сбербанк прогнозирует, что к 2035 году каждое второе рабочее место в России будет связано с использованием систем ИИ.

Однако вместе с этим растёт число этических, социальных и правовых вызовов, связанных с алгоритмической предвзятостью, защитой данных и ответственностью за решения машин.

В ноябре 2024 года Общецерковная аспирантура и докторанттура (ОЦАД) имени святых равноапостольных Кирилла и Мефодия утвердила принципы применения технологий искусственного интеллекта (ИИ) в академических работах своих студентов, преподавателей и научных сотрудников.

В приказе ОЦАД было указано, что вуз одобряет применение ИИ в письменных работах, если оно не нарушает принципы «академической честности» и «прозрачности». При этом генерировать текст для академической работы с помощью ИИ студентам и сотрудникам вуза запрещается. Если же автор написал работу самостоятельно, то он может провести стилистическую обработку текста с помощью таких технологий [1].

В 2015 году компания Google выпустила обновлённое приложение «Google Фото», в котором появилась функция автоматической сортировки фотографий по альбомам на базе технологии компьютерного зрения. Один из пользователей заметил, что приложение пометило его и его девушку тегом «гориллы» [2]. В 2018 году Reuters сообщил, что компания Amazon отказалась от использования нейросети для подбора персонала, так как система демонстрировала гендерную предвзятость и отдавала предпочтение мужчинам [3].

В 2023 году одна из крупнейших в мире технологических компаний Samsung Electronics столкнулась с реальной проблемой, связанной с применением генеративного ИИ в автоматизации рабочих процессов. Сотрудники компании начали использовать чат-боты на базе ИИ, такие как ChatGPT и другие подобные сервисы,

для помощи в решении профессиональных задач: исправлении кода, автоматизации рутин и преобразовании текста.

В ответ на эти инциденты Samsung запретила своим сотрудникам использовать ChatGPT и аналогичные чат-боты на корпоративных устройствах и внутренних сетях без специальных мер защиты данных [4].

Однако это привело к несколько инцидентов утечки конфиденциальной информации. Инженеры Samsung, пытаясь с помощью ChatGPT оптимизировать код или обработать результаты внутренних встреч, вводили в интерфейс ИИ закрытые данные компании, включая прототипы программ, исходные тексты, служебные записи и технологические детали.

В России вопрос этичного применения ИИ в образовательных и оценочных процедурах остаётся в стадии формирования. Министерство науки и высшего образования РФ в публичных заявлениях поддерживает инициативы университетов по разработке этических норм использования ИИ и отмечает необходимость выработки регуляторной и этической рамки. Тем не менее, общие регламенты и процедуры пока отсутствуют.

В частности, при поддержке Правительства РФ в информационном агентстве России 26 октября прошёл международный форум «Этика искусственного интеллекта: начало доверия», где подписании Кодекс этики искусственного интеллекта, разработанного с учетом требований Национальной стратегии развития искусственного интеллекта на период до 2030 года [5].

Для деловой практики в 2025 году Банк России представил Кодекс этики для разработки и использования ИИ на финансовом рынке и сформулировал пять основных принципов его применения [6].

Параллельно в 2023 году в Китае активно развивались проекты внедрения систем искусственного интеллекта в сферу образования. Одним из самых обсуждаемых стал проект так называемых «умных классов» (Smart Classroom) – pilotная инициатива, запущенная в нескольких школах провинций Чжэцзян, Гуандун и Гуйчжоу. Система, разработанная китайскими компаниями Hanwang Technology (Hanvon) и Hikvision, использовала алгоритмы компьютерного зрения для распознавания лиц, фиксации микровыражений и направления взгляда учеников во время урока. На основе этих данных программа вычисляла «индекс вовлечённости» каждого учащегося, определяя, кто слушает внимательно, а кто «отвлёкся».

Китайские школы внедряли такие технологии в рамках национальной стратегии цифрового образования, однако уже через несколько месяцев система вызвала широкий общественный резонанс. Родители и правозащитники заявили, что такая практика нарушает право на частную жизнь детей, а также создает атмосферу постоянного наблюдения и психологического давления.

После публикации критических статей в китайских и зарубежных СМИ проект был приостановлен, а Министерство образования КНР поручило провести аудит подобных ИИ-систем с точки зрения этики и защиты персональных данных.

20 марта 2023 года произошёл инцидент в работе ChatGPT, связанный с утечкой данных из-за ошибки в open-source компоненте redis-py, который использовался сервисом. По объяснению OpenAI, из-за этого сбоя у части пользователей могла

отображаться информация, не относящаяся к их аккаунту – в частности, некоторые пользователи могли видеть названия чужих чатов в боковой панели истории.

OpenAI также сообщила, что у ограниченного числа пользователей ChatGPT Plus в этот период могла быть раскрыта платёжная информация. По оценке компании, инцидент мог затронуть около 1,2% подписчиков ChatGPT Plus, которые были активны в определённый момент времени [7].

Исследование компании Harmonic Security (2025) показало, что около 4 % всех запросов, отправляемых пользователями в ChatGPT и аналогичные сервисы, содержат конфиденциальную корпоративную информацию – пароли, исходный код, персональные данные и внутренние отчёты [8].

Такие действия происходят без злого умысла: сотрудники пытаются ускорить работу, но тем самым создают «теневую ИИ-активность» (shadow AI), не контролируемую отделами информационной безопасности.

Следовательно, нормы регулирования этики применения ИИ в России только зарождаются, глубокого понимания применяя ИИ в разных сферах жизнедеятельности ещё не сформулировано.

Источники:

1. Вуз при РПЦ разрешил применять ИИ в дипломных работах студентов, URL: <https://www.vedomosti.ru/society/articles/2024/11/08/1073949-vuz-razreshil-primenyat>
2. Google распознает на фото чернокожих людей как горилл, URL: https://www.cnews.ru/news/top/google_raspoznaet_na_foto_chernokozhih
3. Insight — Amazon scraps secret AI recruiting tool that showed bias against women, URL: <https://www.reuters.com/article/world/insight-amazon-scaps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G/>
4. Промах Samsung с ChatGPT. Неизбежно?, URL: <https://adguard.com/ru/blog/samsung-chatgpt-leak-privacy.html>
5. В России появился Кодекс этики в сфере искусственного интеллекта, URL: <https://ac.gov.ru/news/page/v-rossii-poavilsa-kodeks-etiki-v-sfere-iskusstvennogo-intellekta-27078>
7. Банк России представил Кодекс этики при использовании искусственного интеллекта, URL: <https://www.garant.ru/news/1829473/>
8. March 20 ChatGPT outage, URL: <https://openai.com/index/march-20-chatgpt-outage/>
9. Workers keep feeding company secrets into ChatGPT, URL: https://www.axios.com/2025/07/31/workers-company-secrets-chatgpt?utm_source=chatgpt.com

Задание:

1. Проанализируйте описанную ситуацию. Используя дополнительные источники информации, выделите ключевые этические, социальные и юридические проблемы, связанные с использованием ИИ.
2. Определите, какие моральные и правовые дилеммы стоят перед разработчиками, пользователями и государством в области этичного применения ИИ.
3. Предложите собственное решение проблем на уровне государства (макроуровень) и компаний (микроуровень). Какие нормы, регламенты или механизмы контроля необходимо ввести, чтобы минимизировать риски использования ИИ?

Требования к решению кейса

Решение кейса должно быть представлено в виде двух файлов:

- 1) Презентация (формат .pdf или .pptx) с основными положениями решения и выводами (не более 15 слайдов);
- 2) Текстовый файл (формат .pdf или .docx) с дополнительной информацией (не более 1 страницы формата А4 12 шрифтом): расчеты, аналитические данные, ссылки на источники информации.

В презентации и текстовом файле должны содержаться разные материалы.

Файлы с решением кейса должны быть отправлены в срок до **16 марта 2026 года 23:59 мск.** по следующему электронному адресу: **innovatika.universiada.hsmi@mail.ru.** Позже решения кейса приниматься не будут.

Основные критерии оценки

При выставлении оценок за решение кейса будут использоваться следующие критерии:

- Качество проведенного анализа и аргументированность сделанных выводов.
- Логика и структура изложения.
- Качество оформления презентации.
- Нестандартность мышления при выработке решения.
- Учет современных экономических особенностей и условий.