

Белая книга этики в сфере искусственного интеллекта в образовании

Научный руководитель – Зуева Анна Сергеевна

Филатов Денис Дмитриевич

Студент (бакалавр)

Московский государственный университет имени М.В.Ломоносова, Высшая школа
государственного аудита, Москва, Россия
E-mail: denf2006den@gmail.com

Сегодня миллионы людей в повседневной жизни пользуются инструментами генеративного ИИ, которые способны имитировать возможности человека по созданию таких материалов, как тексты, изображения, видео, музыка и программные коды [1]. Быстрое развитие ИИ оказывает серьезное влияние на образование, однако адаптация как преподавателей, так и учащихся к таким условиям вызывает проблемы этичного использования искусственного интеллекта. Решение проблемы, как определить, какое поведение при взаимодействии с искусственным интеллектом является этичным, преследует цель сохранить атмосферу доверия и культивировать нормы добросовестности в образовательной среде [1]. Для этого в 2024 году была издана «Белая книга этики в сфере ИИ», в центре внимания которой — 42 ключевых вопроса, которые охватывают широкий спектр тем [2], включая и образование.

В первой главе «Белой книги» рассматриваются десять наиболее значимых проблем в связи с развитием ИИ. Среди них есть те, которые затрагивают учебную деятельность и раскрываются в главе «ИИ в сфере образования». Это проблемы: допустимости создания цифровых имитаций человека, проблема необходимости информирования человека об использовании ИИ, проблема оспаривания решений, принятых с использованием ИИ, проблема предвзятости ИИ, проблема социального рейтингования [2].

Отвечая на вопросы об использовании ИИ в образовании, в «Белой книге» приводятся ряд примеров, когда его применение является этичным и позволяет сделать более эффективным учебный процесс. В частности, отмечается, что преподаватели и учащиеся могут использовать ИИ, но с учетом требований законодательства, возрастных ограничений, особенностей конкретной образовательной сферы и внутреннего положения конкретной организации [2]. Также источник называет этичным написание курсовых и иных учебных работ с помощью ИИ, служащего в качестве инструмента для обработки информации и редактирования текста, если при этом нет подмены авторства. «Белая книга» обращает внимание на применение ИИ в качестве вспомогательного инструмента и при выполнении такой рутинной задачи, как проверка преподавателями ученических работ. Однако окончательное решение всегда должно приниматься преподавателем [2].

«Белая книга» выделяет примеры неэтичного поведения при использовании ИИ в сфере образования. Так, недопустимым отмечается ведение преподавателем предмета через свою цифровую имитацию без личного присутствия в аудитории, поскольку взаимодействие, понимание социально допустимого и этически корректного поведения, навыки работы со смыслами и ценностями могут формироваться только в межличностном общении [2]. Подробно освещается вопрос использования системы прокторинга на основе ИИ, поскольку ее постоянное применение видится неэтичным и нецелесообразным. Тотальный контроль подрывает доверие между учащимися и педагогами, нарушает психологически комфортную атмосферу, необходимую для эффективного обучения [2]. Стоит отметить и вопрос допустимости снижения оценки учащемуся при подозрении на использование ИИ в учебе. Такой подход кажется авторам «Белой книги» необоснованно категоричным [2].

Положения, закрепленные в «Белой книге» носят рекомендательный характер. Вместе с тем, согласно Руководству по использованию генеративного искусственного интеллекта в образовании и научных исследованиях ЮНЕСКО [1], чтобы разрешить противоречия вокруг генеративного ИИ и использовать потенциальные преимущества генеративного ИИ в образовании, его необходимо регулировать. Регулирование генеративного ИИ в образовательных целях требует ряда шагов и политических мер, основанных на человеко-ориентированном подходе, для обеспечения его этического, безопасного, справедливого и значимого использования. В Руководстве ЮНЕСКО предлагаются шаги для регулирования, одним из которых является утверждение и внедрение конкретных правил этики ИИ. Обзор существующих национальных стратегий ИИ, проведенный ЮНЕСКО в 2023 году, показывает, что выявление таких этических проблем и формулирование руководящих принципов является общим только для примерно 40 национальных стратегий ИИ. Требуется внедрить этические принципы в действующие законы или правила, и это происходит редко. Только около 20 стран определили какие-либо четкие правила этики ИИ, в том числе в отношении образования, либо в рамках национальных стратегий ИИ, либо иным образом [1]. Чтобы использование искусственного интеллекта в образовании не стало препятствием для прогресса, необходимо правовое регулирование. Для этого в ряде стран разрабатываются и принимаются нормативные акты в сфере ИИ. Так, для стран ЕС принят Закон об ИИ (AI Act) [3]. Документ призван защитить от высокорискового искусственного интеллекта (ИИ) гражданские права, стимулировать инновации и вывести Европу в лидеры в области ИИ. В конце октября 2023 г. президент США Джо Байден подписал Указ о безопасном, надежном и заслуживающем доверия ИИ. Данный документ требует от создателей систем ИИ прозрачности процессов: они обязаны делиться с американским правительством данными о результатах тестирования безопасности и другой критически важной информацией до того, как разработка выйдет на рынок [4]. Также в марте 2024 года в КНР был опубликован проект закона «Об искусственном интеллекте». В проекте провозглашается принцип развития сферы ИИ в сочетании с государственным контролем за отраслью, а также принцип соответствия отрасли ИИ принятым в Китае мерам по этике [5]. В 2019 году Указом Президента РФ была принята Национальная стратегия развития искусственного интеллекта на период до 2030 года, одним из принципов которой является разработка и внедрение образовательных модулей в рамках образовательных программ всех уровней образования, программ повышения квалификации и профессиональной переподготовки [6]. Сравнивая правовое регулирование ИИ в разных странах, можно проследить общую тенденцию: искусственный интеллект развивается быстрее, чем принимаются акты, регулирующие его. Поэтому этические нормы заменяют правовые, становясь фундаментом для будущего законодательного регулирования.

В заключение стоит отметить, что «Белая книга» закрепляет этические нормы, которые служат ориентиром для взаимодействия преподавателей и обучающихся с ИИ и впоследствии призваны стать фундаментом для правового регулирования данных технологий в образовании. В современном мире, где алгоритмы могут написать текст или решить задачу, главной ценностью становится не результат, а этический путь его достижения, направление которого показывает Белая книга этики в сфере ИИ.

Источники и литература

- 1) Руководство по использованию генеративного искусственного интеллекта в образовании и научных исследованиях. С. 1-27 <https://aspnet-unesco.ru/prod/files/doc/event/c4d192913ca89aa65b274ac99c2b4214.pdf>
- 2) Белая книга этики в сфере искусственного интеллекта / Под ред. А. В. Незнамова. М., 2024. С. 2-133.

- 3) AI Act <https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai>
- 4) ИСИЭЗ <https://issek.hse.ru/news/886273942.html>
- 5) Саяпин С.П. О правовом регулировании генеративного искусственного интеллекта в Китае // Право и политика. 2025. № 3. С. 19-29.
- 6) Указ Президента РФ от 10.10.2019 N 490 (ред. от 15.02.2024) "О развитии искусственного интеллекта в Российской Федерации