

СОЦИАЛЬНО-ГУМАНИТАРНАЯ ЭКСПЕРТИЗА СИСТЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ОБРАЗОВАНИИ: ОБОСНОВАНИЕ, СУЩНОСТЬ, РЕКОМЕНДАЦИИ ПО ПРОВЕДЕНИЮ

Е. В. Брызгалина, Е. М. Шкомова

Московский государственный университет имени М. В. Ломоносова

Поступила в редакцию 12 февраля 2023 г.

Аннотация: в статье представлены рекомендации по проведению социально-гуманитарной экспертизы применения проектов ИИ в образовании. Описаны возможные цели и принципы социально-гуманитарной экспертизы ИИ в образовании, а также значимые характеристики экспертных групп. Рекомендации для экспертизы проектов ИИ в образовании включают учет полисубъектности каждого этапа реализации проекта; выявление обоснованности применения ИИ в образовании; ориентировку на цели и задачи образования; учет ценностно-нормативного характера образования; необходимость сохранности персональных данных; внимание к статусу несовершеннолетних пользователей системы ИИ; учет динамики индивидуальной и групповой информации в процессах сбора, передачи и хранения данных; учет рисков дискриминации и монополизации образования; предотвращение рисков нарушения коммуникации; информирование всех субъектов образования о работе ИИ.

Ключевые слова: философия образования, искусственный интеллект в образовании, этика искусственного интеллекта, социально-гуманитарная экспертиза.

Abstract: the article presents recommendations for conducting a social and humanitarian expertise of AI projects in education. Possible goals and principles of social and humanitarian expertise of AI in education are described, as well as significant characteristics of expert groups evaluating the applicability of AI. Recommendations for the expertise of AI projects in education include taking into account the involvement of many subjects at all stages of project implementation; identifying the validity of the use of AI in education; orientation to the goals and objectives of education; taking into account the value and normative nature of education; the need for the safety of personal data; registration of minor users of the AI system; accounting for individual and group information in the processes of collecting, transmitting and storing data; taking into account risks of discrimination and monopolization of education; prevention of communication disruption risks; informing subjects of education about AI.

Key words: Philosophy of Education, Educational AI, Ethics of AI, Social and Humanitarian Expertise.

Образование является сферой расширяющегося применения систем искусственного интеллекта (СИИ). В системе российского образования процесс внедрения ИИ в образование происходит в рамках проекта цифровизации образования [1]. СИИ могут использоваться для индивидуализации и персонификации процесса образования; для алгоритмизации оценивания результатов образования; для обеспечения технологического принятия решений в образовательной политике [2]. Реализация мер по достижению этих целей за счет использования технологий ИИ в образовании сопровождается рядом этических проблем и угроз, связанных с высокой рискогенностью технологий ИИ. Необходима социально-гуманитарная и этическая экспертиза применения СИИ в образовании.

В настоящий момент в России не существует законодательно закрепленных форм социально-гуманитарной или этической экспертизы использования технологий ИИ в образовании. Необходима нормативная проработка вопросов организации социально-гуманитарной экспертизы, ее задач, субъектов осуществления и форм ее проведения. Рекомендации, представленные в данной статье, основаны на обзоре основных проблем, рисков и ограничений использования ИИ в образовании – как тех, которые уже зафиксированы, так и тех, которые прогнозируются отечественными и зарубежными авторами в научных публикациях 2019–2022 гг. [3–4]. Данные предложения нуждаются в обсуждении со стороны экспертного сообщества для возможного закрепления их в формате принимаемых сообществом разработчиков и пользователей рекомендаций для социально-гума-

© Брызгалина Е. В., Шкомова Е. М., 2023



Контент доступен под лицензией Creative Commons Attribution 4.0 License.
The content is available under Creative Commons Attribution 4.0 License.

нитарной экспертизы проектов применения ИИ в образовании.

Специфика социально-гуманитарной экспертизы ИИ в образовании

Социально-гуманитарная экспертиза – это не подлежащая окончательной алгоритмизации исследовательская практика, постоянно функционирующая на всех этапах получения знания и использования технологий в конкретных проектах [5]. Она ориентирована на выявление и оценку гуманитарной и социальной приемлемости поставленных целей, использованных методов, полученных результатов проектной деятельности. Концептуальные и методологические основания социально-гуманитарной экспертизы (и этической экспертизы, как ее разновидности) были заложены в рамках отечественной биоэтической традиции Б. Г. Юдиным и успешно развиваются по настоящее время в отношении рискогенных проектов [6]. Для масштабирования концепции социально-гуманитарной экспертизы в сфере образования как одной из рискогенных сфер необходима синхронизация общих принципов этики, этики ИИ и специальных принципов для ИИ в области образования [7]. Общие этические принципы ИИ должны быть более точно определены и детализированы, должны предполагать ограничения, обусловленные спецификой образовательной сферы. Цель социально-гуманитарной экспертизы проектов ИИ в образовании должна заключаться в обеспечении защиты человеческого достоинства, соблюдении и защите прав субъекта, чьи интересы затрагивает конкретный проект СИИ.

В настоящий момент не существует общепризнанного пула субъектов, который может быть задействован в проведении социально-гуманитарной экспертизы применительно к проектам применения ИИ в образовании. С учетом полисубъектности полного цикла реализации проектов ИИ в образовании возможно включение в состав группы экспертизы представителей различных заинтересованных групп: экспертов в техническом создании ИИ, анализе данных и глубоком обучении; теоретиков и практиков образования; специалистов по этике искусственного интеллекта; представителей родительской общественности; представителей групп обучающихся; независимых представителей общественности. Внутри отдельного проекта социально-гуманитарную и этическую экспертизу может проводить специально организованная экспертная группа (департамент/комитет). В формате «мини-исследования» экспертизу могут проводить отдельные эксперты и представители общественности.

При проведении социально-гуманитарной и этической экспертизы экспертные группы могут руководствоваться существующими документами госу-

дарственного и надгосударственного уровней, направленные на защиту здоровья и жизни человека, его прав и свобод, а также на документы, закрепляющие этические принципы применения ИИ:

Конституция Российской Федерации (принята всенародным голосованием 12.12.1993 г. с изменениями, одобренными в ходе общероссийского голосования 01.07.2020 г.);

Федеральный закон от 21.11.2011 г. № 323-ФЗ (в ред. от 28.12.2022 г., с изм. и доп., вступ. в силу с 01.03.2023 г.) «Об основах охраны здоровья граждан в Российской Федерации»;

Федеральный закон от 24.07.1998 г. № 124-ФЗ (в ред. от 03.04.2023 г.) «Об основных гарантиях прав ребенка в Российской Федерации»;

Федеральный закон от 29.12.2012 г. № 273-ФЗ (в ред. от 17.02.2023 г.) «Об образовании в Российской Федерации»;

Федеральный закон от 27.07.2006 г. № 152-ФЗ «О персональных данных»;

Федеральный закон от 27.07.2006 г. № 149-ФЗ «Об информации, информационных технологиях и о защите информации»;

Федеральный закон от 27.12.2002 г. № 184-ФЗ «О техническом регулировании»;

Федеральный закон от 29.12.2010 г. № 436-ФЗ «О защите детей от информации, причиняющей вред их здоровью и развитию»;

Указ Президента РФ от 10.10.2019 г. № 490 «О развитии искусственного интеллекта в Российской Федерации» (вместе с Национальной стратегией развития искусственного интеллекта на период до 2030 года);

Указ Президента РФ от 09.05.2017 г. № 203 «О Стратегии развития информационного общества в Российской Федерации на 2017–2030 годы»;

Решение Комиссии Таможенного союза от 28.05.2010 г. № 299 «О применении санитарных мер в Евразийском экономическом союзе»;

Резолюция 2015/2103 (INL) Civil Law Rules on Robotics;

доклады экспертов Всемирной комиссии по этике научных знаний и технологий ЮНЕСКО (доклады об этике робототехники (2017), об этике ИИ (2019)); «Азиломарские принципы ИИ» (2017);

Монреальская декларация об ответственном развитии искусственного интеллекта (2017);

Руководство по этике для надежного ИИ Специальной группы экспертов высокого уровня Совета Европы (2018);

Глобальная инициатива этики автономных и интеллектуальных систем (IEEE, 2016);

Рекомендуемые практики по управлению качеством наборов данных для медицинского искусственного интеллекта ((IEEE);

Модельная конвенция робототехники и искусственного интеллекта (2018);

Концепция развития регулирования отношений в сфере технологий искусственного интеллекта и робототехники до 2024 года (2020);

Кодекс этики искусственного интеллекта (2021).

Пул документов, в которых излагаются руководящие принципы и указания по этике, постоянно пополняется. Эта деятельность в настоящее время активно происходит в нашей стране в связи с процессами отказа от автоматического распространения на российские практики рекомендаций международных организаций. Приоритет сохранения национальной суверенности образования дополнительно актуализирует задачи выработки рекомендаций по этической экспертизе СИИ для целей образования.

Рекомендации по проведению социально-гуманитарной экспертизы проектов ИИ в образовании

С учетом перспектив, проблем, рисков и ограничений применения ИИ в образовании в работе групп социально-гуманитарной экспертизы технологий ИИ в образовании целесообразно учитывать следующие рекомендации:

1. Учитывать при проведении экспертизы, что на каждом этапе реализации проекта ИИ в образовании участвуют множество субъектов, как внешних по отношению к образованию, так и непосредственно причастных к процессу образования и фиксации его результатов.

Необходимо учитывать, что на каждом этапе разработки, внедрения и использования ИИ в процесс включены различные субъекты, представляющие различные целевые и ценностные позиции по отношению к проекту: заказчики, технические разработчики, субъекты управления системой образования, администрация образовательного учреждения, методисты, преподаватели и учащиеся и их родители. Междисциплинарное сотрудничество, доступность информации, возможность давать и получать обратную связь о работе системы, — все это необходимо для продвижения к доверенному и прозрачному ИИ [8–9]. Для выявления социально-этических проблем применения СИИ необходимо создавать площадки для коммуникации задействованных субъектов [10].

2. При проведении экспертизы выявлять обоснованность разработки данной системы, учитывая риски применения данной системы ИИ в образовании (в том числе риски, которые могут быть вызваны неправильными решениями, предложенными СИИ).

Для проведения экспертизы на этапе подготовки к реализации проекта должна быть представлена цель создания СИИ, а также научное обоснование и описание ее практической ценности и актуальности. Обоснование создания данной системы должно быть подкреплено цитированием научных публикаций.

При машинном обучении должны быть использованы только качественные наборы данных, а процесс их создания и подготовки должен быть прозрачен для эксперта, проводящего этическую оценку СИИ. В технической документации, предоставляемой производителем должны быть предоставлены описания источников информации, на основе которых были сформированы наборы данных для обучения и тестирования; условий, в которых формировались исходные данные; процедуры его создания, включая критерии включения/выключения, размерности наборов данных, методы анализа (Data collection/Study Data). Важно убедиться, что СИИ обучалась на той же популяции, для использования в которой она разрабатывается. В итоговом описании конкретной СИИ должны быть представлены окончательные характеристики включенных в наборы данных обучающихся (базовые демографические признаки, доступные предикторы), в том числе количество лиц (случаев) с отсутствующими, неполными данными и результатом. В качестве характеристик набора данных обязательно должны быть указаны: общий объем; классы, баланс классов, количество записей в каждом классе; наличие и вид разметки данных; переменные-предикторы; стандарты предоставления данных.

Значительное повышение доверия к СИИ обеспечивает проведение производителем независимой внешней валидации моделей на наборах данных, которые не были доступны производителю и не были использованы для обучения, как максимум — валидация независимой группой экспертов.

3. При проведении экспертизы ориентироваться на цели и задачи образования, нормативно установленные для уровня и форм образования, для которых разрабатывается конкретный продукт ИИ.

Для проверки того, соответствует ли предназначение каждого продукта ИИ и последствия его использования целям и задачам образования, необходимо участие в разработке СИИ субъектов, представляющих интересы образования [8]. При применении ИИ для управления системой образования, необходимо учитывать, что цели и задачи образования являются качественными и не могут сводиться исключительно к количественным параметрам, с которыми оперируют СИИ поддержки экспертных решений при администрировании образования [11–12].

Для того, чтобы продукт соответствовал различным образовательным ситуациям, а его функционал отвечал различным задачам, исследователи предлагают на этапе разработки системы закладывать множественные подходы в его работе [13].

4. Ориентироваться при проведении экспертизы на ценностно-нормативный характер сферы образования.

При реализации проекта образовательного ИИ необходимо учитывать специфику образования как

особой сферы социальных отношений, имеющих ценностно-нормативное содержание. Ключевые риски применения ИИ в образовании связаны с задачами сохранения экзистенциальной безопасности человека и антропологической сущности образования [2; 14; 15]. Необходимо сформулировать ценностные и нормативные принципы применения ИИ для защиты автономии человека, его психического и морального благополучия. Необходима проверка того, насколько цифровые помощники готовы давать адекватные ответы на различные запросы пользователей с точки зрения моральных норм [16–17]. Также необходимо, чтобы работа СИИ предполагала постоянное обновление входных данных, для того чтобы пользователь получал релевантные запросу и актуальные данные [18]. Необходимо способствовать созидательным целям применения ИИ в образовании в соответствии с ценностно-нормативными ориентирами.

5. Обеспечивать сохранность личных данных, зафиксированных в образовательном процессе в силу их высокой чувствительности для обучающихся субъектов и семей обучающихся.

Персонализация и индивидуализация образования при помощи СИИ предполагает повышение точности профилирования пользователей и сбор чувствительной личной информации о пользователе, в том числе биометрических данных [19]. Необходима институционализация и унификация процедуры защиты данных, а также создание процедуры проверки того, как компании, предоставляющие продукты ИИ, обеспечивают конфиденциальность данных пользователей ИИ [20].

6. Учитывать, что пользователями и конечными бенефициарами систем на основе ИИ могут быть несовершеннолетние учащиеся.

Необходимо учитывать, что конечным пользователем СИИ может быть несовершеннолетний учащийся. В таком случае в принципах работы СИИ должна быть учтена ограниченность автономии принятия решений несовершеннолетним субъектом, невозможность получения добровольного информированного согласия, особенности уровня психологического развития личности, несформированность ценностных оснований личности, необходимость сохранности личных данных несовершеннолетних пользователей. В связи с цифровым мониторингом и профилированием несовершеннолетних учеников возникает проблема «открытого будущего» – сохранения свободы человека с течением времени отказываться от участия в цифровых процессах.

7. Учитывать в разработке принципов сбора, передачи и хранения данных на основе ИИ специфичное для образовательных практик сочетание индивидуальной и групповой информации.

Для сферы образования характерно то, что в ситуации обучения сочетаются индивидуальные и групповые данные [21]. Необходима разработка принципов хранения и передачи данных, обладающих различными характеристиками сроков хранения и степени открытости.

8. Учитывать риски алгоритмической дискриминации или социальной дискриминации, вызванной использованием ИИ, а также риски «монополизации» образования.

Одна из проблем применения ИИ в образовании касается дискриминации пользователей, выраженной в туннельности предлагаемых алгоритмами образовательных траекторий и поисковой информации [22]. Решение данной проблемы требует ручного управления системой и сторонней оценки результатов ее работы. Также дискриминация может возникать на социальном уровне, когда пользователь подвергается социальным ограничениям из-за опоры исключительно на данные ИИ или из-за утечки личных данных, собранных алгоритмами.

При проведении экспертизы необходимо учитывать, что монополизация образовательного контента и доступа к системам ИИ в руках одного субъекта – государства или отдельных образовательных компаний чревата высокими рисками [23].

9. Выявлять и предотвращать риски нарушения коммуникации в сфере образования.

Посредничество ИИ трансформирует социальные связи, важные для процессов обучения и воспитания. Следует учитывать выводы исследователей о том, что традиционные субъекты образования не должны заменяться системами ИИ, но должны приобретать новые характеристики, необходимые в условиях нарастающей цифровизации [17; 24]. Таким образом, необходимо способствовать сохранению полноты человеческой коммуникации, значения роли преподавателя как наставника и роли специалиста в корректировке алгоритмических функций системы. Нельзя допустить алгоритмизации учебного процесса и исключения человека как активного деятеля из социальных процессов [15].

10. Способствовать информированию субъектов образования о целях, возможностях, рисках и ограничениях применения систем ИИ в образовании.

Подобно добровольному информированному согласию в медицине информирование относительно применения ИИ в образовании, проблемах, ограничениях и рисках должно быть заблаговременным и исчерпывающим для данного уровня развития СИИ. Чем конкретнее ожидания пользователей от СИИ, тем меньше риски злоупотребления их доверием [13].

Необходимость введения этической экспертизы как механизма регулирования полного цикла разработки и применения систем ИИ в образовании обу-

словлена тем, что практики применения ИИ вызывают проблемы и риски, касающиеся антропологической сущности образования, измерения социальной жизни и личной жизни отдельного человека. Риск-генность применения технологий ИИ в образовании усугубляется тем, что сфера образования представляет собой область массового применения ИИ. Социально-гуманитарную экспертизу проектов ИИ в образовании на каждом этапе реализации проекта необходимо разрабатывать превентивно, компенсируя и предвзятым самым отставание правового регулирования прикладного ИИ, применяемого в отдельных отраслях практики. Проблемы и риски разработки и внедрения систем ИИ в образование требуют создания комплексной системы этико-правового регулирования, предусматривающей формирование общих и отраслевых правовых норм, а также функционирования социально-гуманитарной экспертизы по сопровождению проектов.

ЛИТЕРАТУРА

1. Паспорт стратегии Цифровая трансформация образования / Министерство просвещения Российской Федерации. – 15.07.2021. – PURL: <https://docs.edu.gov.ru/document/267a55edc9394c4fd7db31026f68f2dd/>
2. Брызгалова Е. В. Искусственный интеллект в образовании. Анализ целей внедрения / Е. В. Брызгалова // Человек. – 2021. – Т. 32, № 2. – С. 1.
3. Брызгалова Е. В. Ключевые проблемы, риски и ограничения применения ИИ в медицине и образовании / Е. В. Брызгалова, А. Н. Гумарова, Е. М. Шкомова // Вестник Московского университета. Сер. 7, Философия. – 2022. – № 6 – С. 93–108.
4. Брызгалова Е. В. Философский анализ ограничений использования систем искусственного интеллекта в образовании / Е. В. Брызгалова // Интеллектуальные системы : теория и приложения. – 2022. – Т. 26, № 1. – С. 367–370.
5. Брызгалова Е. В. Роль социально-гуманитарной экспертизы в обеспечении научного лидерства Российской Федерации / Е. В. Брызгалова, В. Н. Киселев // The Digital Scholar : Philosopher's Lab = Цифровой ученый: лаборатория философа. – 2021. – Т. 4, № 3. – С. 44–46.
6. Юдин Б. Г. От этической экспертизы к экспертизе гуманитарной / Б. Г. Юдин // Гуманитарное знание : тенденции развития в XXI веке. В честь 70-летия Игоря Михайловича Ильинского / колл. моногр. ; под ред. В. А. Лукова. – М. : Изд-во Нац. ин-та бизнеса, 2006. – С. 214–237.
7. Latham A. A survey of the general public's views on the ethics of using AI in education. / A. Latham, S. Goltz // S. Isotani, E. Millán, A. Ogan, P. Hastings, B. McLaren, R. Luckin (Eds.). Artificial Intelligence in Education : 20th International Conference, AIED 2019, Chicago, IL, USA, June 25–29, 2019, Proceedings, Part I, P. 194–206.
8. Ethics of AI in Education : Towards a Community-Wide Framework / W. Holmes, K. Porayska-Pomsta, K. Holstein [et. al] // International Journal of Artificial Intelligence in Education. – 2022. – № 32. – P. 504–526.
9. Chaudhry M. A Transparency Index Framework for AI in Education / M. Chaudhry, M. Cukurova, R. Luckin. – URL: https://www.researchgate.net/publication/360732021_A_Transparency_Index_Framework_for_AI_in_Education
10. Kitto K. Practical ethics for building learning analytics / K. Kitto, S. Knight // British Journal of Educational Technology. – 2019. – № 50.
11. Shum B. S. J. Learning analytics and AI : Politics, pedagogy and practices / B. S. J. Shum, R. Luckin // British Journal of Educational Technologies. – 2019. – № 50. – P. 2785–2793.
12. Цветкова И. В. Симулякры цифрового гуманизма в современном образовании / И. В. Цветкова // Вестник Воронеж. гос. ун-та. Сер.: Философия. – 2021. – № 2. – С. 88–97.
13. Ruane E. Conversational AI : Social and Ethical Considerations / E. Ruane, A. Birhane, A. Ventresque. – URL: https://ceur-ws.org/Vol-2563/aics_12.pdf
14. Согомонов А. Ю. Цифровой университет в информационном обществе : этические аспекты текущей трансформации / А. Ю. Согомонов // Ведомости прикладной этики. – 2021. – № 57. – С. 48–63.
15. Искусственный интеллект в исследованиях сознания и общественной жизни (к 70-летию статьи А. Тьюринга «Вычислительные машины и разум») : материалы круглого стола / В. А. Лекторский, Е. А. Алексеева, Н. Н. Емельянова [и др.] // Философия науки и техники. – 2022. – Т. 27. – № 1. – С. 5–33.
16. Chinedu W. O. Chatbots applications in education : a systematic review / W. O. Chinedu, A. Abejide // Computers and Education : Artificial Intelligence. – 2021. – Vol. 2.
17. Richards D. Supporting and challenging learners through pedagogical agents : Addressing ethical issues through designing for values / D. Richards, V. Dignum // British Journal of Educational Technologies. – 2019. – № 50. – P. 2885–2901.
18. Cunningham-Nelson S. A review of chatbots in education : Practical steps forward. / S. Cunningham-Nelson, W. Boles, L. Trouton, E. Margerison // 30th Annual Conference for the Australasian Association for Engineering Education (AAEE 2019) : Educators Becoming Agents of Change : Innovate, Integrate, Motivate. – Engineers Australia, Australia. – P. 299–306.
19. Williamson B. New digital laboratories of experimental knowledge production : Artificial intelligence and education research / B. Williamson // London Review of Education. – 2020. – № 18 (2). – P. 209–220.
20. Leaton Gray S. Artificial intelligence in schools : towards a democratic future / S. Leaton Gray // London Review of Education. – 2020. – № 18 (2). – P. 163–177.
21. Pammer-Schindler V. Data-Related Ethics Issues in Technologies for Informal Professional Learning / V. Pammer-Schindler, C. Rosé // International Journal of Artificial Intelligence in Education. – 2021. – № 32. – P. 609–635.

22. *Johansen J.* Studying human-to-computer bias transference. / J. Johansen, T. Pedersen, C. Johansen // AI & Soc. – 2021. – URL: <https://arxiv.org/pdf/2005.08231.pdf>

23. *Согомонов А. Ю.* Цифровая этика для цифрового образования в цифровом мире / А. Ю. Согомонов // Ведомости прикладной этики. – 2021. – № 58. – С. 17–30.

Московский государственный университет имени М. В. Ломоносова

Брызгалина Е. В., кандидат философских наук, доцент

E-mail: evbrz@yandex.ru

Шкомова Е. М., кандидат философских наук, ассистент кафедры этики

E-mail: eshkomova@yandex.ru

24. *Yang S. J. H.* Human-centered artificial intelligence in education : Seeing the invisible through the visible / S. J. H. Yang, H. Ogata, T. Matsui, N. Chen // Computers and Education : Artificial Intelligence. – 2021. – № 2.

Moscow State University named after M. V. Lomonosov
Bryzgalina E. V., PhD in Philosophy, Associate Professor

E-mail: evbrz@yandex.ru

Shkomova E. M., PhD in Philosophy, Assistant of the Department of Ethics

E-mail: eshkomova@yandex.ru