

УДК 378

ЭТИЧЕСКИЕ АСПЕКТЫ ВНЕДРЕНИЯ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА В ОБРАЗОВАТЕЛЬНЫЕ СИСТЕМЫ: АНАЛИЗ ПРОБЛЕМ И ПУТИ РЕШЕНИЯ

КОРЧАГИН ПАВЕЛ АНАТОЛЬЕВИЧстарший преподаватель
ФГАОУ ВО "Казанский (Приволжский) федеральный университет**КОРЧАГИН ИЛЬЯ ПАВЛОВИЧ**Руководитель направления,
АНО ДПО "Корпоративный университет Сбербанка"
Академия технологий, данных и кибербезопасности

Аннотация: В статье рассматриваются ключевые этические проблемы, возникающие при интеграции систем искусственного интеллекта (ИИ) в образовательный процесс. Проводится анализ таких аспектов, как предвзятость и дискриминация, конфиденциальность и защита данных, автономия и самостоятельность учащихся, а также прозрачность и подотчетность ИИ-систем. На основе проведенного анализа предлагаются возможные пути решения выявленных этических проблем.

Ключевые слова: искусственный интеллект, этика, предвзятость, конфиденциальность, автономия, прозрачность, подотчетность.

ETHICAL ASPECTS OF IMPLEMENTING ARTIFICIAL INTELLIGENCE IN EDUCATIONAL SYSTEMS: ANALYSING PROBLEMS AND SOLUTIONS

**Korchagin Pavel Anatolievich,
Korchagin Ilya Pavlovich**

Abstract: This article examines the key ethical issues that arise when integrating artificial intelligence (AI) systems into the educational process. Such aspects as bias and discrimination, privacy and data protection, learner autonomy and autonomy, as well as transparency and accountability of AI systems are analysed. Based on the analyses, possible solutions to the identified ethical issues are proposed.

Keywords: artificial intelligence, ethics, bias, privacy, autonomy, transparency, accountability.

Введение. Стремительное развитие технологий искусственного интеллекта открывает новые горизонты для совершенствования образовательных систем. Потенциал ИИ в повышении качества и эффективности обучения не вызывает сомнений [1]. Однако внедрение столь мощного инструмента в сферу образования сопряжено с рядом серьезных этических вызовов, требующих тщательного рассмотрения.

Целью данного исследования является анализ ключевых этических проблем, возникающих при интеграции ИИ в образовательный процесс, и поиск возможных путей их решения.

Каждая из проблем анализируется с точки зрения ее потенциального влияния на образовательный процесс, а также предлагаются возможные стратегии по ее преодолению.

Предвзятость и дискриминация. Одной из наиболее острых этических проблем, связанных с внедрением ИИ в образовательные системы, является риск усиления существующих предубеждений и дискриминации. Эта проблема коренится в самой природе обучения ИИ-систем на основе имеющихся данных.

Источники предвзятости. Алгоритмы машинного обучения, лежащие в основе современных ИИ-систем, обучаются на больших массивах данных, которые могут содержать скрытые предубеждения и стереотипы, присущие обществу. Например, если исторические данные об успеваемости студентов отражают существовавшие в прошлом гендерные или расовые предубеждения, ИИ-система может неосознанно воспроизводить эти паттерны в своих рекомендациях или оценках.

Последствия алгоритмической предвзятости. Предвзятость в ИИ-системах может привести к серьезным негативным последствиям в образовательном процессе:

- несправедливая оценка способностей учащихся из недостаточно представленных групп;
- ограничение образовательных возможностей для определенных категорий студентов;
- усиление существующего социального неравенства в доступе к качественному образованию;
- формирование у учащихся искаженного представления о своих способностях и потенциале.

Пути решения проблемы предвзятости. Для минимизации рисков, связанных с алгоритмической предвзятостью, необходимо предпринять следующие шаги:

- обеспечение репрезентативности данных: при сборе и подготовке обучающих данных для ИИ-систем необходимо уделять особое внимание включению информации о различных демографических группах;
- разработка алгоритмов, снижающих предвзятость: использование техник машинного обучения, специально разработанных для минимизации влияния скрытых предубеждений в данных;
- междисциплинарный подход: привлечение специалистов из различных областей (социологов, психологов, педагогов) к разработке и тестированию ИИ-систем для образования;
- регулярный аудит и мониторинг: внедрение систем постоянного контроля за результатами работы ИИ для выявления и устранения возможной предвзятости;
- повышение осведомленности: обучение разработчиков, педагогов и учащихся распознаванию и противодействию алгоритмической предвзятости.

Конфиденциальность и защита данных. Внедрение ИИ в образовательные системы неизбежно связано с обработкой значительных объемов персональных данных учащихся и преподавателей, что поднимает серьезные вопросы о конфиденциальности и безопасности этой информации.

Риски, связанные с обработкой персональных данных. Использование ИИ в образовании предполагает сбор и анализ разнообразной информации о студентах, включая:

- академические показатели;
- поведенческие паттерны;
- личные предпочтения и интересы;
- социально-экономический статус;
- медицинские данные.

Неправомерное использование или утечка такой информации могут иметь серьезные последствия для учащихся, включая нарушение приватности, дискриминацию и потенциальные злоупотребления.

Проблемы обеспечения конфиденциальности данных в образовательных ИИ-системах включают несколько значительных вызовов. Масштаб данных представляет собой одну из ключевых проблем, так как ИИ-системы требуют огромных объемов информации для эффективной работы, что увеличивает риски утечек [2]. Долгосрочное хранение данных также является фактором риска, поскольку длительное хранение образовательной информации повышает вероятность ее компрометации.

Интеграция различных образовательных платформ и систем создает дополнительные уязвимости, так как взаимодействие между этими системами может облегчить несанкционированный доступ к данным. Недостаточная осведомленность пользователей о том, какие данные собираются и как они

используются, усиливает проблему, поскольку многие не осознают масштаба и потенциальных рисков, связанных с обработкой их личной информации.

Стратегии защиты данных и обеспечения конфиденциальности. Для минимизации рисков, связанных с обработкой персональных данных в образовательных ИИ-системах, необходимо реализовать следующие меры:

- строгие протоколы безопасности: внедрение современных методов шифрования и защиты данных на всех этапах их сбора, хранения и обработки;
- анонимизация: использование технических решений для обезличивания данных без потери их аналитической ценности;
- принцип минимизации данных: сбор и хранение только тех данных, которые действительно необходимы для функционирования образовательных ИИ-систем;
- прозрачность и информированное согласие: предоставление учащимся и их родителям полной информации о собираемых данных и целях их использования;
- регулярные аудиты безопасности: проведение независимых проверок систем безопасности и практик обработки данных;
- образовательные программы: повышение цифровой грамотности и осведомленности о вопросах приватности среди всех участников образовательного процесса;
- нормативно-правовое регулирование: разработка и внедрение строгих правовых норм, регламентирующих обработку персональных данных в образовательных ИИ-системах.

Автономность и самостоятельность учащихся. Внедрение ИИ в образовательный процесс поднимает важные вопросы о сохранении автономности и развитии самостоятельности учащихся. Существует риск, что чрезмерная зависимость от ИИ-систем может ограничить способность студентов к критическому мышлению и самостоятельному принятию решений.

Потенциальные угрозы автономности учащихся включают несколько ключевых аспектов. Алгоритмическое ограничение выбора может возникнуть, когда ИИ-системы предлагают ограниченный набор образовательных возможностей, основываясь на предыдущем поведении студентов. Это может препятствовать исследованию широкого спектра учебных материалов и образовательных путей.

Чрезмерная зависимость от рекомендаций ИИ представляет собой еще одну проблему, так как учащиеся могут начать полагаться на технологии для принятия решений, что негативно скажется на развитии их навыков самостоятельного мышления и критического анализа. Такая зависимость ограничивает способность студентов самостоятельно оценивать и выбирать образовательные ресурсы.

Стандартизация образовательных траекторий также является значительной угрозой. Несмотря на возможности персонализации, ИИ может способствовать унификации образовательных путей, что приводит к снижению разнообразия опыта и ограничению перспектив студентов. Это может мешать развитию уникальных навыков и индивидуальных интересов.

Чрезмерное использование ИИ может также снизить роль человеческого взаимодействия, уменьшая важность живого общения между учащимися и преподавателями. Снижение человеческого фактора может негативно повлиять на развитие социальных навыков и эмоционального интеллекта, что является важной составляющей образовательного процесса.

Стратегии сохранения и развития автономности учащихся. Для обеспечения баланса между преимуществами искусственного интеллекта и сохранением автономности учащихся предложены следующие подходы. Развитие критического мышления является ключевым элементом, достигаемым через внедрение курсов, направленных на развитие навыков критического анализа информации и принятия решений. Это помогает студентам осознанно использовать ИИ и критически оценивать его рекомендации.

Обеспечение прозрачности алгоритмов также играет важную роль. Учащимся необходимо понимать, как функционируют ИИ-системы, используемые в их обучении, чтобы они могли взаимодействовать с ними осознанно и уверенно.

Расширение возможностей выбора включает предоставление учащимся широкого спектра образовательных опций, не ограничиваясь только рекомендациями ИИ. Это позволяет развивать самостоятельность и ответственность за их собственное обучение.

Поощрение самостоятельности является еще одним важным аспектом. Это может быть достигнуто через разработку заданий и проектов, требующих независимого исследования и принятия решений, что способствует развитию исследовательских навыков и уверенности в своих возможностях.

Сохранение баланса между ИИ и человеческим взаимодействием также является критическим. Несмотря на использование технологий, важно сохранять значительную роль преподавателей в образовательном процессе для обеспечения менторской поддержки и развития социальных навыков у студентов.

Персонализация обучения с учетом целей учащихся предполагает разработку ИИ-систем, которые учитывают не только текущие показатели, но и долгосрочные цели и устремления студентов [3]. Это способствует более точной и индивидуализированной поддержке их учебного процесса.

Включение курсов по этическим аспектам использования ИИ в образовательные программы необходимо для того, чтобы учащиеся могли критически оценивать влияние технологий на свою жизнь. Обучение этике ИИ помогает студентам осознавать и обсуждать потенциальные риски и преимущества использования этих технологий, что способствует формированию ответственного и осознанного подхода к их применению.

Прозрачность и объяснимость ИИ-систем. Одной из ключевых этических проблем при использовании ИИ в образовании является обеспечение прозрачности и объяснимости работы алгоритмов [4]. Непрозрачность процессов принятия решений ИИ-системами может подорвать доверие к ним и затруднить их эффективное использование в образовательном контексте.

Проблемы, связанные с непрозрачностью ИИ-систем:

1. "Черный ящик": многие современные ИИ-системы, особенно основанные на глубоком обучении, функционируют как "черные ящики", что затрудняет понимание логики их решений.
2. Сложность для неспециалистов: технические аспекты работы ИИ могут быть непонятны для большинства участников образовательного процесса.
3. Потенциальные ошибки: отсутствие прозрачности затрудняет выявление и исправление ошибок в работе ИИ-систем.
4. Этические дилеммы: непрозрачность алгоритмов может скрывать этически спорные решения или предубеждения.

Стратегии повышения прозрачности и объяснимости ИИ в образовании. Для решения проблемы непрозрачности ИИ-систем в образовании предлагается несколько подходов. Один из них включает разработку интерпретируемых моделей, которые отличаются структурной прозрачностью и легкостью интерпретации. Это позволяет лучше понять, как именно принимаются решения ИИ.

Также важным подходом является внедрение методов объяснимого ИИ, которые позволяют получать понятные и доступные объяснения решений, принимаемых ИИ-системами. Для того чтобы пользователи могли интуитивно понять ход рассуждений ИИ, предлагается разработка графических интерфейсов, визуализирующих процессы принятия решений.

Образовательные инициативы играют ключевую роль в повышении осведомленности и компетентности среди преподавателей, студентов и администраторов. Обучающие программы помогут лучше понять основы работы ИИ-систем и их влияние на образовательный процесс.

Открытый доступ к информации является еще одним важным аспектом. Предоставление заинтересованным сторонам доступа к документации, описывающей принципы работы ИИ-систем в образовательных учреждениях, способствует повышению прозрачности и информированности.

Регулярные аудиты алгоритмов также необходимы для проверки соответствия ИИ-систем заявленным целям и этическим нормам. Независимые проверки помогают выявлять и устранять потенциальные проблемы.

Наконец, вовлечение пользователей через механизмы обратной связи позволяет учащимся и преподавателям сообщать о любых возникающих проблемах или неясностях. Это способствует улучшению работы ИИ-систем и повышению их доверия среди пользователей.

Последствия повышения прозрачности ИИ в образовании. Повышение прозрачности и объяснимости ИИ-систем в образовательной сфере может привести к ряду положительных изменений.

Прежде всего, это способствует укреплению доверия к технологиям, так как понимание принципов работы ИИ повышает уверенность всех участников образовательного процесса в справедливости и надежности этих систем. Кроме того, возможность анализа и корректировки работы ИИ-систем позволяет существенно улучшить качество принимаемых решений, что ведет к более точному и эффективному образовательному процессу.

Одновременно с этим, изучение основ работы ИИ способствует развитию критического мышления у учащихся. Понимание того, как функционируют ИИ-системы, помогает им выработать навыки критического анализа и осмысления технологических решений.

Прозрачность в работе ИИ способствует этической согласованности, обеспечивая соответствие функционирования этих систем установленным этическим нормам и ценностям образовательных учреждений. Это позволяет избежать конфликта интересов и поддерживает высокие стандарты профессиональной и моральной ответственности.

Подотчетность и ответственность. Внедрение ИИ-систем в образование поднимает сложные вопросы о подотчетности и ответственности за принимаемые решения и их последствия. Определение четких рамок ответственности является критически важным для этичного использования ИИ в образовательном процессе.

Проблемы распределения ответственности. Проблемы распределения ответственности в процессе внедрения и использования ИИ в образовании охватывают множество аспектов. Вовлечены различные стороны - разработчики, образовательные учреждения, преподаватели, учащиеся. Способность ИИ принимать самостоятельные решения усложняет вопрос об ответственности за эти решения. Кроме того, сложность прогнозирования всех возможных результатов работы ИИ-систем в динамичной образовательной среде создает непредвиденные последствия. Неопределенность в сфере законодательства также является значительной проблемой, поскольку существующее законодательство часто не успевает за развитием технологий ИИ.

Стратегии обеспечения подотчетности и ответственности. Для решения проблем, связанных с подотчетностью и ответственностью при использовании ИИ в образовании, предлагается ряд подходов. Прежде всего, необходима разработка этических кодексов, создающих четкие руководящие принципы для разработки и применения ИИ в образовательных контекстах.

Важно также определить зоны ответственности, четко разграничив ее между разработчиками, образовательными учреждениями и конечными пользователями.

Создание механизмов надзора, таких как независимые комиссии или советы для мониторинга использования ИИ в образовании, является еще одним ключевым аспектом.

Наряду с этим, следует внедрить системы аудита, обеспечивающие регулярные проверки ИИ-систем на соответствие этическим нормам и заявленным целям.

Разработка правовых рамок через сотрудничество с законодателями необходима для создания нормативно-правовой базы, регулирующей использование ИИ в образовании. Не менее важно разработать механизмы страхования рисков от потенциальных негативных последствий использования ИИ.

Наконец, обеспечение прозрачности и объяснимости ИИ-систем позволит проследить и объяснить процесс принятия решений, что критически важно для поддержания доверия к этим технологиям в образовательной сфере.

Распределение ответственности между участниками. Для эффективного распределения ответственности среди участников процесса предлагается следующая модель. Разработчики искусственного интеллекта несут ответственность за техническую надежность и безопасность систем, а также за их соответствие установленным этическим принципам. Образовательные учреждения обязаны осуществлять выбор и внедрение ИИ-систем, обучать персонал и мониторить их использование. Преподаватели ответственны за интеграцию ИИ в учебный процесс, интерпретацию полученных результатов и поддержание баланса между ИИ и традиционными методами обучения. Учащиеся и родители должны использовать ИИ-системы корректно, предоставлять обратную связь и соблюдать этические нормы при работе с ИИ. Регулирующие органы отвечают за разработку и обновление нормативно-правовой базы, а также за контроль соблюдения этических стандартов.

Заключение и рекомендации. Интеграция искусственного интеллекта в образовательные системы представляет собой сложный и многогранный процесс, сопряженный с рядом существенных этических вызовов. Проведенный анализ выявил ключевые проблемные области и наметил потенциальные пути их решения.

Обобщение ключевых этических проблем:

1. Предвзятость и дискриминация: риск усиления существующих социальных неравенств через алгоритмическую предвзятость.
2. Конфиденциальность и защита данных: необходимость обеспечения безопасности личной информации учащихся и преподавателей.
3. Автономия и самостоятельность учащихся: сохранение баланса между поддержкой ИИ и развитием критического мышления.
4. Прозрачность и объяснимость ИИ-систем: обеспечение понимания принципов работы ИИ всеми участниками образовательного процесса.
5. Подотчетность и ответственность: четкое распределение ответственности за решения и действия, связанные с использованием ИИ.

Комплексные рекомендации по этичному внедрению ИИ в образование. На основе проведенного анализа предлагаются следующие рекомендации:

1. Разработка этических рамок:
 - создание универсальных этических принципов для разработки и применения ИИ в образовании;
 - регулярное обновление этических кодексов с учетом развития технологий и общественных ценностей.
2. Междисциплинарный подход:
 - привлечение специалистов из различных областей (педагогика, психология, этика, право, информатика) для комплексной оценки и разработки ИИ-систем;
 - создание междисциплинарных исследовательских групп для изучения долгосрочных последствий использования ИИ в образовании.
3. Повышение цифровой грамотности:
 - внедрение образовательных программ по ИИ и этике для учащихся, преподавателей и администраторов;
 - развитие навыков критического анализа информации и алгоритмических решений.
4. Усиление защиты данных:
 - внедрение передовых технологий шифрования и анонимизации данных;
 - разработка строгих протоколов доступа к персональным данным учащихся.
5. Обеспечение прозрачности:
 - создание понятных интерфейсов для объяснения решений ИИ-систем;
 - регулярное информирование всех заинтересованных сторон о принципах работы и обновлениях ИИ-систем.
6. Развитие нормативно-правовой базы:
 - сотрудничество с законодателями для создания адекватной правовой основы использования ИИ в образовании;
 - разработка механизмов правовой защиты в случае нарушений или ошибок ИИ-систем.
7. Непрерывный мониторинг и оценка:
 - внедрение систем регулярного аудита ИИ на предмет этичности и эффективности;
 - создание механизмов обратной связи для всех участников образовательного процесса.
8. Поддержка автономности учащихся:
 - разработка ИИ-систем, поощряющих самостоятельное мышление и принятие решений;
 - сохранение баланса между персонализацией обучения и разнообразием образовательного опыта.

Этичное внедрение ИИ в образовательные системы требует комплексного, междисциплинарного подхода и постоянного внимания к возникающим этическим вызовам. Реализация предложенных рекомендаций может способствовать созданию образовательной среды, в которой ИИ служит инструментом поддержки и расширения возможностей, а не источником этических проблем.

Важно подчеркнуть, что этичное внедрение ИИ в образование — это не единовременное действие, а непрерывный процесс, требующий постоянного внимания, рефлексии и адаптации. По мере развития технологий ИИ будут возникать новые этические вызовы, которые потребуют своевременного выявления и решения.

Будущие исследования в этой области должны быть направлены на разработку конкретных методологий оценки этических рисков ИИ в образовании, создание стандартов этического проектирования образовательных ИИ-систем, а также на изучение долгосрочных последствий использования ИИ в образовательном процессе.

Только при условии тщательного учета этических аспектов и ответственного подхода к внедрению, искусственный интеллект сможет полностью реализовать свой потенциал в сфере образования, способствуя созданию более эффективной, справедливой и инклюзивной образовательной среды.

Список источников

1. Амиров Р.А., Билалова У.М. Перспективы внедрения технологий искусственного интеллекта в сфере высшего образования. Управленческое консультирование. 2020;(3):80-88. <https://doi.org/10.22394/1726-1139-2020-3-80-88> (дата обращения: 01.08.2024)
2. Antoniak, M. (2023, June 22). Using large language models with care - AI2 blog. Medium. <https://blog.allenai.org/using-large-language-models-with-care-eeb17b0aed27> (дата обращения: 01.08.2024).
3. Чулюков В. А., Дубов В. М. Искусственный интеллект и будущее образования // Современное педагогическое образование. 2020. №3. С. 27-31.
4. Ethical Considerations for Educational AI // School Education Gateway. URL: <https://school-education.ec.europa.eu/en/discover/news/ethical-considerations-educational-ai> (дата обращения: 01.08.2024).