

Психологические аспекты цифровизации образования / Psychological Aspects of the Digitalization of Education

Научная статья

Original article

DOI: [10.12731/2658-4034-2026-17-1-1096](https://doi.org/10.12731/2658-4034-2026-17-1-1096)

EDN: [NNOBTU](https://nnobtu.ru)

УДК 378.147



Этика и психологическая безопасность применения алгоритмов искусственного интеллекта в системе оценки знаний и профориентации обучающихся

М.В. Сыротюк

*Государственный морской университет имени адмирала Ф.Ф. Ушакова,
Новороссийск, Российская Федерация*

Аннотация

Обоснование. Статья посвящена педагогическому анализу этических и психологических аспектов безопасного использования алгоритмов искусственного интеллекта (ИИ) в двух, критически значимых для образовательной траектории обучающихся, практиках: оценивании учебных достижений и профориентационном сопровождении. В работе обосновано, что переход к алгоритмическим решениям меняет логику образовательного взаимодействия, создавая риски редукции личности обучающегося и утраты его субъектности.

В исследовании выделены четыре базовых параметра контроля алгоритмических систем: педагогическая применимость, социальная справедливость, объяснимость и поддержка автономии обучающегося. Особое внимание уделено операционным индикаторам психологической безопасности, таким как воспринимаемая справедливость, наличие процедур апелляции и недирективность профориентационных рекомендаций.

Практическая значимость исследования заключается в разработке системы рекомендаций для высших учебных заведений по созданию этико-педагогической инфраструктуры, необходимой для трансформации ИИ из инструмента внешнего контроля в средство поддержки развития и самоопределения обучающихся.

Цель исследования состоит в разработке и обосновании модели педагогических условий, снижающих риски нарушения психологической безопасности и обеспечивающих этически ответственное применение алгоритмических решений в университетской среде.

Материалы и методы. Методологическую основу составили системный и деятельностный подходы, компетентностная перспектива оценки, а также принципы педагогической этики и психологии безопасности образовательной среды. Иссле-

дование реализовано в форме теоретико-аналитического проекта: выполнен критический обзор ключевых направлений исследований, а именно, алгоритмическая справедливость, применимость и надежность цифровой оценки, объяснимость моделей, риск-ориентированное управление. Проведено концептуальное моделирование и предложены отдельные индикаторы психологической безопасности при использовании ИИ в образовательных решениях.

Результаты. Ключевым результатом является проект «этической и психологически безопасной алгоритмизации», включающий три уровня (технологический, педагогико-организационный, субъектно-коммуникативный) и четыре контура контроля (применимость, справедливость, прозрачность, поддержка автономии обучающегося). Показано, что наиболее значимые риски локализуются не в самом факте применения ИИ, а в педагогическом «обрамлении» алгоритма: в способах интерпретации результата, распределении ответственности, режимах обратной связи и возможностях апелляции.

Ключевые слова: алгоритмическая оценка; психологическая безопасность; педагогическая этика; справедливость; недискриминация; профориентация обучающихся; объяснимость моделей; ответственность за решение

Для цитирования. Сыротюк, М. В. (2026). Этика и психологическая безопасность применения алгоритмов искусственного интеллекта в системе оценки знаний и профориентации обучающихся. *Russian Journal of Education and Psychology*, 17(1-1), 587–607. <https://doi.org/10.12731/2658-4034-2026-17-1-1096>

Ethics and psychological safety of using artificial intelligence algorithms in the knowledge assessment and career guidance system for students

M.V. Sirotiyk

*Maritime State University of Admiral F.F. Ushakov,
Novorossiysk, Russian Federation*

Abstract

Background. The article is devoted to the pedagogical analysis of the ethical and psychological aspects of the safe use of artificial intelligence (AI) algorithms in two practices that are critically important for the educational trajectory of students: assessment of academic achievements and career guidance. The paper substantiates that the transition to algorithmic solutions changes the logic of educational interaction, creating risks of reducing the student's personality and the loss of his subjectivity.

The study identifies four basic control parameters of algorithmic systems: pedagogical applicability, social justice, explainability, and support for student autonomy.

Special attention is paid to operational indicators of psychological safety, such as perceived fairness, the availability of appeal procedures, and the non-directivity of career guidance recommendations. The practical significance of the research lies in the development of a system of recommendations for higher education institutions to create the ethical and pedagogical infrastructure necessary for the transformation of AI from an instrument of external control into a means of supporting the development and self-determination of students.

Purpose. The purpose of the research is to develop and substantiate a model of pedagogical conditions that reduce the risks of psychological safety violations and ensure the ethically responsible application of algorithmic solutions in the university environment.

Materials and methods. The methodological basis is based on systematic and activity-based approaches, a competence-based assessment perspective, as well as the principles of pedagogical ethics and the psychology of educational environment safety. The research was implemented in the form of a theoretical and analytical project: a critical review of key research areas was carried out, namely, algorithmic fairness, applicability and reliability of digital assessment, explainability of models, risk-based management. Conceptual modeling has been carried out and separate indicators of psychological safety have been proposed when using AI in educational solutions.

Results. The key result is the project of «ethical and psychologically safe algorithmization», which includes three levels (technological, pedagogical-organizational, subject-communicative) and four control circuits (applicability, fairness, transparency, support for student autonomy). It is shown that the most significant risks are localized not in the very fact of using AI, but in the pedagogical «framing» of the algorithm: in the ways of interpreting the result, the distribution of responsibility, feedback modes and appeal possibilities.

Keywords: algorithmic assessment; psychological safety; pedagogical ethics; fairness; non-discrimination; student career guidance; explainability of models; responsibility for decision

For citation. Sirotiyk, M. V. (2026). Ethics and psychological safety of using artificial intelligence algorithms in the knowledge assessment and career guidance system for students. *Russian Journal of Education and Psychology*, 17(1-1), 587–607. <https://doi.org/10.12731/2658-4034-2026-17-1-1096>

Введение

Актуальность темы определяется тем, что алгоритмы ИИ в высшей школе становятся инфраструктурой принятия решений: они участвуют в автоматизации тестирования, анализе учебной аналитики, прогнозировании успеваемости, выявлении «групп риска», а также в профориентации – от рекомендаций образовательных треков до соотнесения

компетенций с рынком труда. Для педагогической науки данная тенденция значима не как технологическая модернизация сама по себе, а как изменение логики образовательного взаимодействия: оценка и профориентация переходят из преимущественно диалогических практик в частично алгоритмизованные процедуры, где возрастает риск редукции личности к набору признаков, смещения целей обучения к «оптимизации показателей» и перераспределения авторитета от педагога и студента к системе рекомендаций [2, с. 20].

Постановка проблемы состоит в противоречии между выделяемыми преимуществами ИИ: оперативность обратной связи, масштабирование оценочных процедур, персонализация образовательных маршрутов, и потенциальными деформациями образовательной среды: снижением психологической безопасности, усилением неравенства через скрытую дискриминацию, утратой субъектности студента и размыванием ответственности за педагогически значимое решение. Проблема обостряется тем, что в оценивании и профориентации ошибка алгоритма имеет не только «измерительную», но и жизненно-траекторную стоимость: неверная оценка влияет на академический статус, самоуважение и мотивацию, в то время, как неверная рекомендация – на выбор профессии и возможности самореализации [4, с. 71].

При исследовании вопросов этики и психологической безопасности применения алгоритмов искусственного интеллекта в системе оценки знаний и профориентации обучающихся целесообразно сформулировать следующую гипотезу: если внедрение алгоритмов ИИ в оценивании и профориентации будет организовано как педагогически управляемая система, а не как техническая автоматизация, включающая прозрачность критериев, процедуру апелляции, объяснимую обратную связь и недирективное консультирование, то риски психологической небезопасности, будь то тревожность, стигматизация, внешняя атрибуция успеха/неуспеха, снижение автономии, будут статистически и качественно снижаться по сравнению с ситуацией, где ИИ используется как «черный ящик» и источник окончательного вердикта [14, с. 627].

В педагогическом аспекте оценивание рассматривается как деятельность, выполняющая диагностическую, мотивационную, регулятивную

и развивающую функции. В этой рамке качество оценки определяется не только точностью измерения, но и педагогической применимостью (валидностью): соответствием образовательным целям, понятностью критериев, поддержкой саморегуляции и развитием учебной автономии [11, с. 61].

Цифровая оценка усилила измерительную компоненту, но одновременно поставила вопрос о соотношении количественных индикаторов и качественных характеристик достижений. Когда к цифровой оценке добавляется ИИ, меняется статус оценочного решения: оно приобретает вид «объективизированного» вывода системы, что в сознании обучающихся и иногда преподавателей повышает авторитет результата независимо от реальной «валидности» модели. Этот феномен можно описать как эффект алгоритмического авторитета: склонность принимать алгоритмический вывод как более нейтральный и точный по сравнению с человеческой оценкой, даже при наличии сомнений в исходных данных и критериях [10, с. 944].

В профориентации ключевой является идея субъектности выбора и поддержка профессионального самоопределения. Современные профориентационные модели опираются на принцип недирективности, учет личностных смыслов, развитие рефлексии и проектирование индивидуальной траектории [1, с. 922]. Алгоритмические рекомендации потенциально полезны как средство расширения информационного поля и выявления скрытых закономерностей в компетенциях и интересах, однако они несут риск подмены консультации рекомендацией «по умолчанию». В этом случае профориентация может трансформироваться из развивающей практики в распределительный механизм: студенту предлагается «оптимальный» путь, который воспринимается как предписание. В педагогике это трактуется как сдвиг от сопровождения к управлению, что конфликтует с целью формирования самостоятельности [16, с. 49].

Этический контур применения ИИ обычно описывается через принципы справедливости, недискриминации, прозрачности, ответственности, защиту данных и подотчетность. Для педагогики важно уточнение: этика здесь не ограничивается соблюдением формальных норм, а вклю-

чает гуманитарный смысл образования, признающий ценность личности, право на ошибку и возможность развития. Иначе говоря, алгоритм может быть юридически корректен, но педагогически «травматичен», если подает результат в форме окончательного ярлыка или лишает обучающегося пространства для роста [6, с. 32].

Психологическая безопасность образовательной среды традиционно связывается с субъективным ощущением защищенности, возможностью задавать вопросы, признавать трудности без страха унижения, а также с предсказуемостью и справедливостью требований [5, с. 280]. В контексте алгоритмической оценки и профориентации психологическая безопасность становится зависящей от трех факторов. Во-первых, от интерпретируемости критериев: если обучающийся не понимает, почему получен данный результат, растут тревожность и ощущение беспомощности. Во-вторых, от управляемости ситуации: без процедуры апелляции и альтернативной оценки возникает переживание фатализма. В-третьих, от коммуникативного сопровождения: даже корректный алгоритм может стать источником стресса при дефиците уважительной обратной связи и поддержки [7, с. 301].

В исследованиях по алгоритмической справедливости выделяются разные типы перекоса: несбалансированность данных, исторические и социальные асимметрии, смещение метрик качества, эффект прокси-переменных, когда модель использует признаки, внешне нейтральные, но косвенно отражающие социальные различия [15, с. 176]. Для образования это имеет особую значимость: успех обучающегося обусловлен не только способностями, но и доступом к ресурсам, особенностями учебной среды, языковыми и культурными практиками.

Обзор также обнаруживает противоречие между стандартизацией и индивидуализацией. Алгоритмические системы часто оптимизируются под единый показатель точности, тогда как педагогическая реальность требует многомерности: учет прогресса, усилий, условий обучения, динамики мотивации. Возникает пробел: в научных и прикладных работах нередко обсуждается техническая корректность ИИ, но недостаточно разработаны педагогические регламенты его включения в образовательное взаимодействие - как сообщать результаты, как

согласовывать их с педагогическим наблюдением, как обеспечивать развивающую функцию оценки и консультирования [9, с. 85].

Материалы и методы

Исследование выполнено в парадигме теоретико-методологического анализа с элементами концептуального моделирования и проектирования.

Методологическую основу составили:

- Системный подход, позволяющий рассматривать алгоритмы ИИ как подсистему образовательной организации, взаимодействующую с педагогическими, организационными и коммуникативными компонентами;
- Деятельностный подход, фокусирующийся на анализе оценивания знаний и профориентации как видов педагогической деятельности с выделением функциональных блоков сбора информации, интерпретации, принятия решений и обратной связи;
- Компетентностная перспектива, ориентированная на результаты образования, где качество алгоритмических решений оценивается через вклад в формирование и развитие профессиональных, метапредметных и личностных компетенций обучающихся.

В работе исследованы принципы педагогической этики и психологии безопасности использования ИИ, включая постулаты гуманистической педагогики и современной теории психологической безопасности образовательной среды.

Данная работа представляет собой комплексное исследование, выполненное в парадигме теоретико-методологического анализа с элементами концептуального моделирования. Важно подчеркнуть, что это не экспертный опрос или эмпирическое исследование, а теоретико-аналитический проект, включающий критический обзор ключевых направлений современной науки. В ходе работы проведена систематизация рисков, разделенных на инструментальные, институциональные и личностно-психологические кластеры. Основным исследовательским инструментом выступило моделирование, позволившее спроектировать многоуровневую систему «этической и психологически безопасной

алгоритмизации». В рамках методики были разработаны специфические индикаторы безопасности, такие как воспринимаемая справедливость, интерпретируемость критериев и наличие процедур апелляции. Проверка выдвинутой гипотезы осуществлялась на уровне логической обоснованности и концептуальной состоятельности предложенных педагогических условий. Итогом исследования стала разработка системы практических рекомендаций по созданию этико-педагогической инфраструктуры в высшей школе. Таким образом, методика носит теоретико-проектный характер, закладывая необходимую базу для последующей эмпирической апробации предложенных моделей в реальной вузовской практике.

Результаты и обсуждение

В качестве единиц анализа рассмотрены типовые ситуации использования ИИ: автоматизированное оценивание (тесты, задания с закрытым ответом, частично – эссе/проекты через рубрики и ассистирующие модели), предиктивная аналитика (прогноз успеваемости/отчисления), а также профориентационные рекомендации (подбор направлений, компетенций, вакансий, траекторий).

Для каждой ситуации анализировались: образовательная цель, тип решения, уязвимости данных, характер обратной связи, распределение ответственности, возможности студента воздействовать на решение.

Введены четыре группы критериев, позволяющих операционализировать «этическую и психологическую безопасность» применения ИИ:

1. Валидность (применяемость) – как соответствие измеряемой конструкции заявленной цели, доказуемость связи между показателем и компетенцией, наличие педагогической интерпретации;
2. Справедливость – как отсутствие систематических перекосов по социально-демографическим и образовательным подгруппам, недопущение «закрепления низких ожиданий»;
3. Прозрачность и объяснимость, то есть понятность студенту критериев, минимальная достаточная интерпретация результата, фиксация источников данных;

4. Поддержка субъектности посредством возможности апелляции, выбора альтернативных маршрутов, формирующей обратной связи, недирективности профориентационных рекомендаций.

Анализ проводился путем сопоставления требований педагогической этики, задач оценивания и профориентации и типовых архитектур систем с ИИ, а именно: «сбор данных - обработка - вывод - представление результата - управленческое решение». На этой основе конструировались «узлы риска» и проектировались педагогические и организационные меры, минимизирующие негативные последствия.

Результаты исследования могут быть представлены в виде карты ключевых педагогических рисков применения ИИ в оценивании и профориентации, которые целесообразно разделить на три класса: инструментальные, институциональные и личностно-психологические.

В свою очередь, инструментальные риски связаны с использованием некорректных данных и метрик. В оценивании это проявляется в подмене образовательной цели: например, когда тестовая успешность трактуется как полнота компетенции, а качество аргументации - как частотность «правильных» слов. В профориентации инструментальный риск возникает, когда предпочтения и интересы редуцируются к поведенческим следам (клики, выбор курсов) без учета смыслов, а прогноз пригодности строится на неполных данных о личности и контексте [21, p. 442].

Институциональные риски вырастают из организационного внедрения ИИ: неопределенность ответственности, отсутствие процедуры апелляции, давление на преподавателя следовать алгоритмической рекомендации, непрозрачность использования данных, встраивание ИИ в систему санкций вместо системы поддержки. Типовой сценарий – когда прогноз «низкой успешности» запускает ограничительные решения (отчисление, недопуск), а не развивающие меры (тьюторство, изменение стратегии обучения). В профориентации институциональный риск – превращение рекомендаций в механизм распределения студентов по «оптимальным» траекториям исходя из интересов организации, а не самого студента.

Личностно-психологические риски относятся к изменению самовосприятия и мотивации обучающегося. Алгоритмическая оценка спо-

собна усиливать внешнюю атрибуцию: обучающийся объясняет успех или неуспех «решением системы», что снижает чувство контроля и ухудшает саморегуляцию. Кроме того, существуют риски стигматизации - закрепление ярлыка «слабый», «непредрасположенный», «неподходящий» [8, с. 145]. В профориентации это проявляется как эффект самоограничения: обучающийся избегает сложных, но желанных направлений, поскольку алгоритм не «рекомендует» их. Отдельную проблему составляет рост тревожности из-за непрозрачности критериев и невозможности понять, как улучшить результат [17, р. 2150].

Предложенная в исследовании модель «этической и психологически безопасной алгоритмизации» описывает условия ответственного применения ИИ как взаимодействие трех уровней и четырех контуров контроля.

Три уровня:

1. Технологический уровень - качество данных, корректность обучения, мониторинг «отступлений» модели, защита персональных данных, управление доступом. Педагогический смысл данного уровня заключается в том, что «технически верная» модель должна быть применимой (валидной) для образования: измерять то, чему учат, и служить развитию, а не только ранжированию.

2. Педагогико-организационный уровень - регламенты применения, распределение ответственности между преподавателем, администрацией и системой, процедура апелляции, правила представления результата, ограничения на автоматическое принятие решений.

3. Субъектно-коммуникативный уровень - формы обратной связи, обучение студентов пониманию критериев, поддержка рефлексии, диалог о результатах, профилактика стигматизации, развитие цифровой и оценочной грамотности.

Также значимы и четыре контура контроля:

1. Контур применимости (валидности) отвечает на вопрос «что именно измеряется/рекомендуется и почему это педагогически оправдано». Здесь вводится требование конструктивной связности: оценочный индикатор должен быть связан с образовательными результатами и доступен интерпретации через рубрики, примеры, критерии [13, с. 307].

2. Контур справедливости отвечает на вопрос «кому система помогает, а кого - систематически ухудшает в положении». Он требует регулярной проверки различий качества предсказаний по подгруппам, а также анализа ошибок, имеющих жизненно важные последствия.

3. Контур прозрачности отвечает на вопрос «может ли обучающийся понять основания решения и способы улучшения». Прозрачность педагогически трактуется как минимально достаточная объяснимость: не раскрытие технических деталей, а раскрытие критериальной логики, границ применимости и вероятностного характера результатов.

4. Контур поддержки субъектности отвечает на вопрос «сохраняется ли у обучающегося право на выбор и развитие». Включает апелляцию, альтернативные формы оценивания, сопровождение, недирективные рекомендации, стимулирование самооценки и самопланирования.

С педагогической точки зрения психологическая безопасность в условиях алгоритмизации может быть измерена через набор индикаторов, фиксирующих субъективное переживание и объективные организационные условия:

- воспринимаемая справедливость оценивания и рекомендаций;
- понятность критериев и ощущение управляемости (что именно можно сделать, чтобы улучшить результат);
- наличие доверительного канала обсуждения результатов (преподаватель, тьютор, консультант);
- отсутствие уничижительных ярлыков в коммуникации результатов;
- реальная доступность апелляции и пересмотра решения;
- динамика учебной мотивации и вовлеченности после получения алгоритмического результата.

По отношению к исследованиям, фокусирующимся на технической надежности, полученные результаты смещают акцент на педагогическую валидность и коммуникативный дизайн [3, с. 218].

По отношению к работам по этике ИИ модель уточняет специфику образования: образовательная система не может ограничиться принципом «не навреди», поскольку оценка и профориентация должны «развивать», предоставляя пространство для роста и изменения траектории [13, с. 309].

По отношению к психологическим исследованиям безопасности образовательной среды предложено расширение: алгоритмическая прозрачность и управляемость результатов становятся элементами психологической защищенности наряду с межличностным климатом [4, с. 73].

В целом же, интеграция результатов с теоретическими положениями показывает, что основной источник рисков - не алгоритм как вычислительная процедура, а социально-педагогический статус алгоритмического вывода. В образовательной культуре оценка исторически обладает властью: она определяет доступ к ресурсам, признанию, продолжению обучения. Алгоритм, включенный в эту культуру без «педагогических предохранителей», неизбежно усиливает дисциплинарную функцию оценивания и ослабляет развивающую. Аналогично профориентация, будучи смысловой и личностно значимой практикой, при алгоритмизации может стать распределительным механизмом, если рекомендация воспринимается как предписывающий вердикт.

Гипотеза исследования подтверждается на уровне концептуальной обоснованности: при наличии прозрачности критериев, процедур апелляции, формирующей обратной связи и недирективного консультирования логически и психологически уменьшаются основания для тревожности, стигматизации и утраты контроля. Особенно существенной является роль коммуникативного сопровождения: алгоритмический результат следует интерпретировать как сигнал для диалога и поддержки, а не как окончательную оценку личности или «профессиональной пригодности».

Ограничения исследования обусловлены теоретико-проектным характером работы: предложенная модель нуждается в эмпирической апробации на выборках студентов разных направлений и в контекстах различных типов вузов.

Кроме того, результаты не претендуют на универсальное решение для всех видов ИИ-систем: алгоритмы различаются по степени интерпретируемости, зависимостью от контента, качеством данных.

Еще одно ограничение связано с изменчивостью технологий: требования к объяснимости и риски перекоса следует пересматривать по мере обновления моделей и источников данных.

Существенные педагогические проблемы, выявленные в ходе анализа, можно сформулировать следующим образом.

Первая проблема - редукция образовательных целей к измеряемым показателям при автоматизации оценивания, в результате чего «то, что легко измерить» подменяет «то, что важно развивать».

Вторая проблема - размывание ответственности: участники образовательного процесса не понимают, кто отвечает за ошибку и какие процедуры исправления доступны.

Третья проблема - психологическая небезопасность алгоритмической коммуникации результатов: непрозрачный и категоричный вывод провоцирует тревожность, внешнюю атрибуцию и стигматизацию.

Четвертая проблема - директивность профориентационных рекомендаций, когда алгоритм фактически ограничивает пространство профессионального самоопределения.

Теоретические последствия заключаются в необходимости расширить понимание педагогической применимости и значимости цифровой оценки: она должна включать критерии справедливости, прозрачности и поддержки субъектности как интегральные элементы качества. Практические последствия связаны с трансформацией процедур внедрения ИИ в вузе: требуется не только закупка и настройка решения, но и создание этико-педагогической инфраструктуры - регламентов, ролей, компетенций и форм обучения участников [12, с. 124].

Заключение

В статье разработана модель этически и психологически безопасного применения алгоритмов ИИ в оценивании знаний и профориентации обучающихся, основанная на идее педагогически управляемой алгоритмизации.

Достигнута цель исследования: предложена структурированная система уровней и контуров контроля, позволяющая выявлять риски и проектировать условия их снижения.

Задачи исследования выполнены: систематизированы подходы и проблемные зоны, выявлены причинно-следственные механизмы рисков, предложены индикаторы психологической безопасности и сформулированы практические рекомендации.

Перспективы дальнейших исследований включают: эмпирическую проверку предложенных индикаторов безопасности, сравнительный анализ разных моделей объяснимости в образовательных интерфейсах, изучение влияния алгоритмической обратной связи на мотивацию и саморегуляцию, а также разработку методик подготовки преподавателей и тьюторов к работе с алгоритмическими результатами.

Рекомендации и направления преодоления проблем сгруппированы по уровням модели и ориентированы на реализацию в вузовской практике:

1. Педагогическое проектирование алгоритмической оценки и профориентации требует фиксации педагогической цели применения ИИ до выбора инструмента. Так, целесообразно заранее определить:

- для оценивания - какие компоненты компетенции диагностируются и как результат будет использован;
- для профориентации - какую задачу решает рекомендация (расширение вариантов, выявление дефицитов, план развития), а не «подбор единственно правильной профессии».

Также следует запрещать подмену формирующей обратной связи санкционирующими управленческими решениями без участия человека: алгоритм должен инициировать педагогическую поддержку, а не автоматически ограничивать обучающегося и применять принцип «двухканальной оценки» для сложных ситуаций: алгоритмический результат используется как один источник данных и сопоставляется с оценкой преподавателя или комиссии.

2. В рамках этической экспертизы и справедливости - целесообразно ввести регулярную процедуру аудита перекосов, то есть проверять, различается ли точность и характер ошибок у разных подгрупп обучающихся; анализировать «тяжелые ошибки», влияющие на доступ к экзамену, стипендии, перевод или рекомендации траектории. Важно также исключать использование чувствительных данных и их прокси без доказанной педагогической необходимости; в профориентации ограничивать применение данных поведенческой аналитики, если нет ясной связи с профессиональными интересами и компетенциями.

Следует закрепить ответственность, то есть определить, кто утверждает правила использования ИИ, кто принимает окончательное

решение, кто рассматривает апелляции, кто отвечает за корректировку модели и информирование обучающихся.

3. С целью обеспечения прозрачности и объяснимости как педагогической нормы следует:

- разработать «паспорт алгоритмического решения» в образовательной организации: какие данные учитываются, каковы границы применимости, какие типовые ошибки возможны, как обучающийся может улучшить результат или оспорить его;
- переводить объяснение результата в форму педагогически значимой обратной связи: не «вам не подходит», а «по данным текущих достижений и выбранных активностей вероятнее успешность в X; для расширения выбора в Y рекомендуется развить компетенции A, B, C»;
- внедрять минимальные стандарты интерфейса: отображение критериев, диапазонов неопределенности, примеров уровневых ответов (для оценки), а также альтернативных вариантов и условий успеха (для профориентации).

4. Обеспечивая «субъектность» обучающегося и психологическую безопасность, важно гарантировать апелляцию и «право на второй маршрут»: то есть обучающийся должен иметь возможность пересдать, представить альтернативный формат демонстрации компетенции (проект, устный ответ, портфолио), получить консультирование.

Также в рамках данного аспекта следует говорить об оценочной грамотности и цифровой осведомленности обучающихся: как интерпретировать вероятностные прогнозы, почему рекомендации не равны предписаниям, как использовать обратную связь для планирования развития. Важно предотвращать стигматизацию в языке и практиках: исключить категоричные ярлыки, обеспечить конфиденциальность, использовать поддерживающую и развивающую риторику.

Целесообразно выделять роль тьютора/куратора как медиатора между алгоритмом и обучающимся: для обсуждения результата, постановки целей, выбора стратегии.

5. Организационные регламенты внедрения необходимо контролировать посредством создания междисциплинарной комиссии по от-

ветственному применению ИИ в образовании (педагоги, психологи, специалисты по данным, юристы, представители студентов), которая утверждает сценарии использования и критерии приема.

В рамках осуществляемой работы необходимо вводить пилотирование и поэтапное масштабирование: сначала - формирующие сценарии с низким риском, затем - расширение функций при наличии доказательств валидности и безопасности. Также следует контролировать эффект на мотивацию и благополучие обучающихся: включать в оценку эффективности не только «точность прогнозов», но и изменения в вовлеченности, академической добросовестности, уровне тревожности, доверии к оценке.

6. Специфика профорientации должна проявляться через недирективное алгоритмическое консультирование, то есть возможность трактовать алгоритм как инструмент расширения поля возможностей, а не «распределителя». Рекомендации должны включать несколько альтернатив и объяснение условий успешности:

- обеспечивать смысловую работу: обсуждение ценностей, интересов, опыта, жизненных планов, так как алгоритм не заменяет диалог, а дает материал для него.
- фиксировать принцип обратимости: обучающийся вправе пересматривать выбор, а система должна поддерживать динамику самоопределения, а не закреплять ранний прогноз.

В целом же следует сделать вывод, что этически корректное и психологически безопасное применение ИИ в оценке знаний и профорientации возможно лишь при превращении алгоритма из «внешнего судьи» в элемент педагогически организованной поддержки, лишь тогда алгоритмическая рациональность перестает конкурировать с гуманитарной миссией образования и начинает работать на развитие компетенций, автономии и осознанности обучающегося.

Список литературы

1. Алешов, В. В., Самчинская, Я. Б., Алешов, Е. В., Линник, И. И., & Криницкий, Ю. В. (2025). Этические вопросы использования искусственного интеллекта в образовательном пространстве. *Инновационные направления интеграции науки, образования и производства*, с. 922–927. EDN: <https://elibrary.ru/KYHACY>

2. Ахметшина, Э. И., Пономарёв, Н. А., & Абдюшева, А. Д. (2024). Моделирование образовательных траекторий с использованием искусственного интеллекта. *Современная наука: актуальные проблемы теории и практики. Серия: Естественные и технические науки, (1)*, с. 20–23. <https://doi.org/10.37882/2223-2966.2024.01.03>. EDN: <https://elibrary.ru/VDEDYY>
3. Боброва, И. И., & Тимофеев, Е. Г. (2024). Влияние искусственного интеллекта на формирование информационной безопасности в образовании. *Актуальные проблемы современной науки, техники и образования*, с. 218. EDN: <https://elibrary.ru/HLWPBQ>
4. Богатырева, Ю. И., Привалов, А. Н., & Смирнова, В. А. (2025). Модель угроз личности обучающихся при применении систем искусственного интеллекта. *Вестник РГГУ. Серия: Информатика. Информационная безопасность. Математика, (1)*, с. 70–94. <https://doi.org/10.28995/2686-679>. EDN: <https://elibrary.ru/WITTRY>
5. Горбатенко, Н. Ю. (2024). Формирование личности студента в условиях воздействия искусственного интеллекта. *Устойчивое развитие социально-экономических процессов в условиях цифровой трансформации современного общества*, с. 279–282. EDN: <https://elibrary.ru/QOWEFV>
6. Ермакова, Т. А. (2023). Возможности и угрозы применения искусственного интеллекта в современном образовании. *Высшее образование в эпоху глобальных трансформаций: достижения, вызовы, перспективы*, с. 31–35. EDN: <https://elibrary.ru/COEDNB>
7. Козырев, П. С., & Шмакова, А. П. (2025). Этические аспекты применения искусственного интеллекта в оценке образовательных достижений. *Управление качеством образования: проблемы и перспективы*, с. 301–307. EDN: <https://elibrary.ru/WNWJSJ>
8. Наумов, В. Н., & Шубодерова, П. А. (2024). Искусственный интеллект и профессиональный выбор нового поколения. *Менеджмент XXI века: формирование суверенной системы образования Российской Федерации*, с. 145–150.
9. Радугин, А. А., & Радугина, О. А. (2021). Применение искусственного интеллекта в образовательном процессе вуза: технологии, потенциал и проблемы. *Вестник Воронежского государственного университета. Серия: Проблемы высшего образования, (4)*, с. 84–87. EDN: <https://elibrary.ru/KLXUPF>
10. Тамьярова, М. В., Тамьяров, А. В., & Тамьяров, И. А. (2023). Интеллектуальная образовательная среда и этические основы интеллектуальной поддержки современного преподавателя. В *Международный форум Kazan Digital Week — 2023*. Казань, с. 944–949. EDN: <https://elibrary.ru/LHIEGO>
11. Тимченко, В. В. (2024). Обеспечение безопасности применения искусственного интеллекта в образовании: стратегии стандартизации. *Технико-технологические проблемы сервиса, (4)70*, с. 60–65. EDN: <https://elibrary.ru/IMQHNH>

12. Харченко, С. Я., & Гришак, С. Н. (2025). Трансформация образовательного процесса в высшей школе в контексте использования технологий искусственного интеллекта. *Гуманитарная парадигма*, (2)33, с. 124–133. EDN: <https://elibrary.ru/HXULDN>
13. Шлыкова, Н. В. (2024). Оптимизация педагогических технологий в профориентационной деятельности с помощью алгоритмов искусственного интеллекта. В *Профессиональная ориентация и профессиональное самоопределение обучающихся: вызов времени*. Саранск, с. 307–312. EDN: <https://elibrary.ru/ELWCWG>
14. Щербанич, Ф. Е., & Ромашкова, О. Н. Этика использования искусственного интеллекта в образовании. В *Наука, инновации и технологии: от идей к внедрению*, с. 626–628.
15. Яламов, Г. Ю. (2019). Методические подходы к обеспечению информационно-психологической безопасности пользователей интеллектуальных обучающих систем. *Педагогическая информатика*, (4), с. 176–182. EDN: <https://elibrary.ru/WAATZV>
16. Яламов, Г. Ю. (2019). Обеспечение информационно-психологической безопасности пользователей интеллектуальных обучающих систем: методические рекомендации. *Управление образованием: теория и практика*, (4)36, с. 48–59. EDN: <https://elibrary.ru/DJBVSA>
17. Kassymova, G. K., Malinichev, D. M., Lavrinenko, S. V., Panichkina, M. V., Koptyaeva, S. V., & Arpentiev, M. R. (2023). Ethical Problems of Digitalization and Artificial Intelligence in Education: A Global Perspective. *Journal of Pharmaceutical Negative Results*, 14(S2), pp. 2150–2161. <https://doi.org/10.47750/pnr.2023.14.S02.254>. EDN: <https://elibrary.ru/QNZUMK>
18. Kiselev, A. A. (2024). The Harm of Artificial Intelligence for the Education System. В *Theory and Practice of Modern Science: The View of Youth*. Saint Petersburg, pp. 288–292. EDN: <https://elibrary.ru/BOAGTD>
19. Mirtcheva-Ivanova, D. (2024). Application of Artificial Intelligence in E-Learning. *Environment. Technology. Resources*, (2), pp. 208–211. <https://doi.org/10.17770/etr2024vol2.8053>. EDN: <https://elibrary.ru/MFWVLB>
20. Odasheva, F. M. (2022). Possibilities of Using Innovative Educational Technologies to Increase Educational Efficiency. *Theoretical & Applied Science*, (5)109, pp. 769–771. <https://doi.org/10.15863/TAS.2022.05.109.73>. EDN: <https://elibrary.ru/ARHX-MJ>
21. Ponomarev, A., Proletarskiy, A., Serebrennikov, O., Guglya, G., & Korotkevich, O. (2020). The Use of Artificial Intelligence Technologies to Implement a Multi-Factor Analysis of a Student's Actions during an Exam. В *E-Learning and Software for Education Conference*. Bucharest, pp. 442–447. <https://doi.org/10.12753/2066-026>. EDN: <https://elibrary.ru/HEZEJQ>

22. Zinina, O. V., & Olentsova, Ju. A. (2020). Innovative Education Activities, Features of Implementation of the Innovative Process. *Baltic Humanitarian Journal*, 9(3)32, pp. 85–87. <https://doi.org/10.26140/bgz3-2020-0903-0019>. EDN: <https://elibrary.ru/ZYQUIA>

References

1. Aleshov, V. V., Samchinskaya, Ya. B., Aleshov, E. V., Linnik, I. I., & Krinitsky, Yu. V. (2025). Ethical issues of using artificial intelligence in the educational space. In *Innovative Directions of Integration of Science, Education and Production*, 922–927. EDN: <https://elibrary.ru/KYHACY>
2. Akhmetshina, E. I., Ponomarev, N. A., & Abdyusheva, A. D. (2024). Modeling educational trajectories using artificial intelligence. *Modern Science: Actual Problems of Theory and Practice. Series: Natural and Technical Sciences*, (1), 20–23. <https://doi.org/10.37882/2223-2966.2024.01.03>. EDN: <https://elibrary.ru/VDEDYY>
3. Bobrova, I. I., & Timofeev, E. G. (2024). The impact of artificial intelligence on the formation of information security in education. In *Current Problems of Modern Science, Technology and Education*, 218. EDN: <https://elibrary.ru/HLWPBQ>
4. Bogatyreva, Yu. I., Privalov, A. N., & Smirnova, V. A. (2025). Threat model for the personality of students when using artificial intelligence systems. *Bulletin of the Russian State University for the Humanities. Series: Computer Science. Information Security. Mathematics*, (1), 70–94. <https://doi.org/10.28995/2686-679>. EDN: <https://elibrary.ru/WITTRY>
5. Gorbatenko, N. Yu. (2024). Formation of a student's personality under the influence of artificial intelligence. In *Sustainable Development of Socio-Economic Processes in the Context of Digital Transformation of Modern Society*, 279–282. EDN: <https://elibrary.ru/QOWEFV>
6. Ermakova, T. A. (2023). Opportunities and threats of using artificial intelligence in modern education. In *Higher Education in the Era of Global Transformations: Achievements, Challenges, Prospects*, 31–35. EDN: <https://elibrary.ru/COEDNB>
7. Kozyrev, P. S., & Shmakova, A. P. (2025). Ethical aspects of using artificial intelligence in assessing educational achievements. In *Quality Management in Education: Problems and Prospects*, 301–307. EDN: <https://elibrary.ru/WNWJSJ>
8. Naumov, V. N., & Shuboderova, P. A. (2024). Artificial intelligence and professional choice of the new generation. In *Management of the 21st Century: Formation of a Sovereign Education System of the Russian Federation*, 145–150.
9. Radugin, A. A., & Radugina, O. A. (2021). Application of artificial intelligence in the university educational process: Technologies, potential and problems. *Bulletin of Voronezh State University. Series: Problems of Higher Education*, (4), 84–87. EDN: <https://elibrary.ru/KLXUPF>

10. Tam'yarova, M. V., Tam'yarov, A. V., & Tam'yarov, I. A. (2023). Intelligent educational environment and ethical foundations of intellectual support for the modern teacher. In *International Forum Kazan Digital Week — 2023*. Kazan, 944–949. EDN: <https://elibrary.ru/LHIEGO>
11. Timchenko, V. V. (2024). Ensuring the safety of using artificial intelligence in education: Standardization strategies. *Technical and Technological Problems of Service*, (4)70, 60–65. EDN: <https://elibrary.ru/IMQHNH>
12. Kharchenko, S. Ya., & Grishak, S. N. (2025). Transformation of the educational process in higher education in the context of using artificial intelligence technologies. *Humanitarian Paradigm*, (2)33, 124–133. EDN: <https://elibrary.ru/HXULDN>
13. Shlykova, N. V. (2024). Optimization of pedagogical technologies in career guidance activities using artificial intelligence algorithms. In *Career Guidance and Professional Self-Determination of Students: The Challenge of the Time*. Saransk, 307–312. EDN: <https://elibrary.ru/ELWCWG>
14. Shcherbanich, F. E., & Romashkova, O. N. Ethics of using artificial intelligence in education. In *Science, Innovation and Technology: From Ideas to Implementation*, 626–628.
15. Yalamov, G. Yu. (2019). Methodological approaches to ensuring the information and psychological security of users of intelligent learning systems. *Pedagogical Informatics*, (4), 176–182. EDN: <https://elibrary.ru/WAATZV>
16. Yalamov, G. Yu. (2019). Ensuring the information and psychological security of users of intelligent learning systems: Methodological recommendations. *Education Management: Theory and Practice*, (4)36, 48–59. EDN: <https://elibrary.ru/DJBVSA>
17. Kassymova, G. K., Malinichev, D. M., Lavrinenko, S. V., Panichkina, M. V., Kopytyaeva, S. V., & Arpentiev, M. R. (2023). Ethical problems of digitalization and artificial intelligence in education: A global perspective. *Journal of Pharmaceutical Negative Results*, 14(S2), 2150–2161. <https://doi.org/10.47750/pnr.2023.14.S02.254>. EDN: <https://elibrary.ru/QNZUMK>
18. Kiselev, A. A. (2024). The harm of artificial intelligence for the education system. In *Theory and Practice of Modern Science: The View of Youth*. Saint Petersburg, 288–292. EDN: <https://elibrary.ru/BOAGTD>
19. Mirtcheva Ivanova, D. (2024). Application of artificial intelligence in e-learning. *Environment. Technology. Resources*, (2), 208–211. <https://doi.org/10.17770/etr-2024vol2.8053>. EDN: <https://elibrary.ru/MFWVLB>
20. Odasheva, F. M. (2022). Possibilities of using innovative educational technologies to increase educational efficiency. *Theoretical & Applied Science*, (5)109, 769–771. <https://doi.org/10.15863/TAS.2022.05.109.73>. EDN: <https://elibrary.ru/ARHXMJ>
21. Ponomarev, A., Proletarskiy, A., Serebrennikov, O., Guglya, G., & Korotkevich, O. (2020). The use of artificial intelligence technologies to implement a multi-factor

- analysis of a student's actions during an exam. In *E-Learning and Software for Education Conference*. Bucharest, 442–447. <https://doi.org/10.12753/2066-026>. EDN: <https://elibrary.ru/HEZEJQ>
22. Zinina, O. V., & Olentsova, Ju. A. (2020). Innovative education activities, features of implementation of the innovative process. *Baltic Humanitarian Journal*, 9(3)32, 85–87. <https://doi.org/10.26140/bgз3-2020-0903-0019>. EDN: <https://elibrary.ru/ZYQUIA>

ДАННЫЕ ОБ АВТОРЕ

Сыротюк Майя Валерьевна, старший преподаватель кафедры «Гражданское и международное право»
ФГБОУ ВО «Государственный морской университет имени адмирала Ф. Ф. Ушакова»
пр-т Ленина, 93, г. Новороссийск, 353925, Российская Федерация
mayas09@mail.ru

DATA ABOUT THE AUTHOR

Maya V. Sirotyk, Senior Lecturer at the Department of Civil and International Law
Maritime State University of Admiral F.F. Ushakov
93, Lenin Prospect, Novorossiysk, 353925, Russian Federation
mayas09@mail.ru
SPIN-code: 4391-9578
ORCID: <https://orcid.org/0009-0008-4250-0915>
Academia.edu: <https://independent.academia.edu/SirotykM>

Поступила 10.01.2026

После рецензирования 28.01.2026

Принята 03.02.2026

Received 10.01.2026

Revised 28.01.2026

Accepted 03.02.2026