

**Кому:** Министерству цифрового развития, инноваций и аэрокосмической промышленности Республики Казахстан

**ФИО и должность автора:** Убайдулла Д.С., эксперт управления информационных технологий Комитета казначейства Министерства финансов Республики Казахстан

**Дата:** 16.10.2024 г.

## **АНАЛИТИЧЕСКАЯ ЗАПИСКА**

**Тема:** Этика искусственного интеллекта: баланс между инновациями и социальной ответственностью

**Ключевые слова:** Искусственный интеллект, этика, социальная ответственность, регулирование.

### **Введение**

Цель данной аналитической записки — исследовать этические аспекты использования искусственного интеллекта (ИИ) и оценить, как достичь баланса между его инновационным потенциалом и социальной ответственностью. Обоснование исследования связано с необходимостью определения путей минимизации негативных последствий и повышения уровня доверия к ИИ, что становится особенно актуальным в условиях постоянного увеличения его влияния на общество.

Текущая проблема заключается в недостаточной прозрачности работы алгоритмов, а также в потенциальных рисках, связанных с нарушением конфиденциальности, автоматизацией труда и дискриминацией. Методология данного исследования включает анализ научных публикаций, законодательных инициатив и примеров международной практики, что позволяет сформировать комплексный взгляд на ситуацию. В рамках данного исследования будут рассмотрены ключевые этические проблемы, актуальные тенденции, международный опыт и рекомендации по улучшению ситуации.

### **Основная часть**

Существует несколько ключевых факторов, способствующих возникновению этических проблем в сфере ИИ. Во-первых, быстрый прогресс в области технологий часто опережает развитие регуляторных механизмов, создавая пробелы в законодательстве. Алгоритмы ИИ, как правило, разрабатываются без достаточного учета этических норм и последствий для общества [1]. Во-вторых, использование больших данных для обучения ИИ поднимает вопросы о конфиденциальности и

согласии пользователей. Исследования показывают, что многие пользователи не осознают, как их данные могут быть использованы [2].

В-третьих, автоматизация процессов может привести к значительной потере рабочих мест, что требует внимания со стороны государственных органов и бизнеса. Казахстан (Рисунок 2) находится в группе стран с очень высоким индексом развития цифрового правительства (E-Government Development Index) и занимает 24-е место среди 193 стран [7]. Согласно прогнозам, значительная часть профессий может оказаться под угрозой исчезновения в ближайшие десятилетия [3]. Это создает социальное напряжение и требует создания эффективных программ переподготовки для работников.

Table 3.4 Countries leading e-government development in Asia, 2024

Country	Rating class	EGDI rank	Subregion	OSI	HCI	TII	EGDI (2024)	EGDI (2022)
Singapore	VH	3	South-eastern Asia	0.9831	0.9362	0.9881	0.9691	0.9133
Republic of Korea	VH	4	Eastern Asia	1.0000	0.9120	0.9917	0.9679	0.9529
Saudi Arabia	VH	6	Western Asia	0.9899	0.9067	0.9841	0.9602	0.8539
United Arab Emirates	VH	11	Western Asia	0.9163	0.9436	1.0000	0.9533	0.9010
Japan	VH	13	Eastern Asia	0.9427	0.9117	0.9509	0.9351	0.9002
Bahrain	VH	18	Western Asia	0.9030	0.8680	0.9877	0.9196	0.7707
Israel	V3	23	Western Asia	0.8541	0.8739	0.9763	0.9014	0.8885
Kazakhstan	V3	24	Central Asia	0.9390	0.8403	0.9235	0.9009	0.8628

**Рисунок 2 – Страны с очень высоким индексом развития цифрового правительства (E-Government Development Index) [7]**

### Текущие тенденции.

На текущий момент наблюдается рост интереса к разработке этических стандартов для ИИ. Компании, работающие в этой области, начинают осознавать важность соблюдения этических норм, что, в свою очередь, ведет к созданию внутренних комитетов по этике. Тем не менее, многие организации по-прежнему действуют по принципу «первое — прибыль», что создает риски для пользователей и общества в целом [1].

Среди тенденций можно выделить увеличение числа инициатив, направленных на повышение прозрачности работы алгоритмов ИИ. Например, в некоторых странах начали разрабатываться законы, обязывающие компании раскрывать информацию о своих алгоритмах и методах обработки данных [4]. Это может привести к повышению доверия к технологиям и снижению уровня общественного недовольства.

### Возможные последствия

Если не будет найден баланс между инновациями и этическими нормами, последствия могут быть серьезными: от юридических исков до общественного недовольства. Отсутствие доверия к ИИ может затормозить его внедрение и развитие, а также угрожать экономическому росту. В то же время, правильный подход к разработке и внедрению ИИ может привести к значительным преимуществам: повышению производительности и улучшению качества жизни [5].

Отсутствие адекватного регулирования может также привести к распространению предвзятостей в алгоритмах, что создаст дополнительные проблемы для отдельных групп населения. Поэтому важно развивать диалог между разработчиками технологий, государственными органами и обществом.

### **Вызовы и возможности.**

Основные вызовы включают необходимость адаптации законодательства и регулирования под новые реалии. В то же время, это открывает возможности для создания новых рабочих мест в сфере разработки этических стандартов и законодательства. Сценарии развития включают как позитивные (разработка международных стандартов), так и негативные (возрастание недовольства и протестов против технологий).

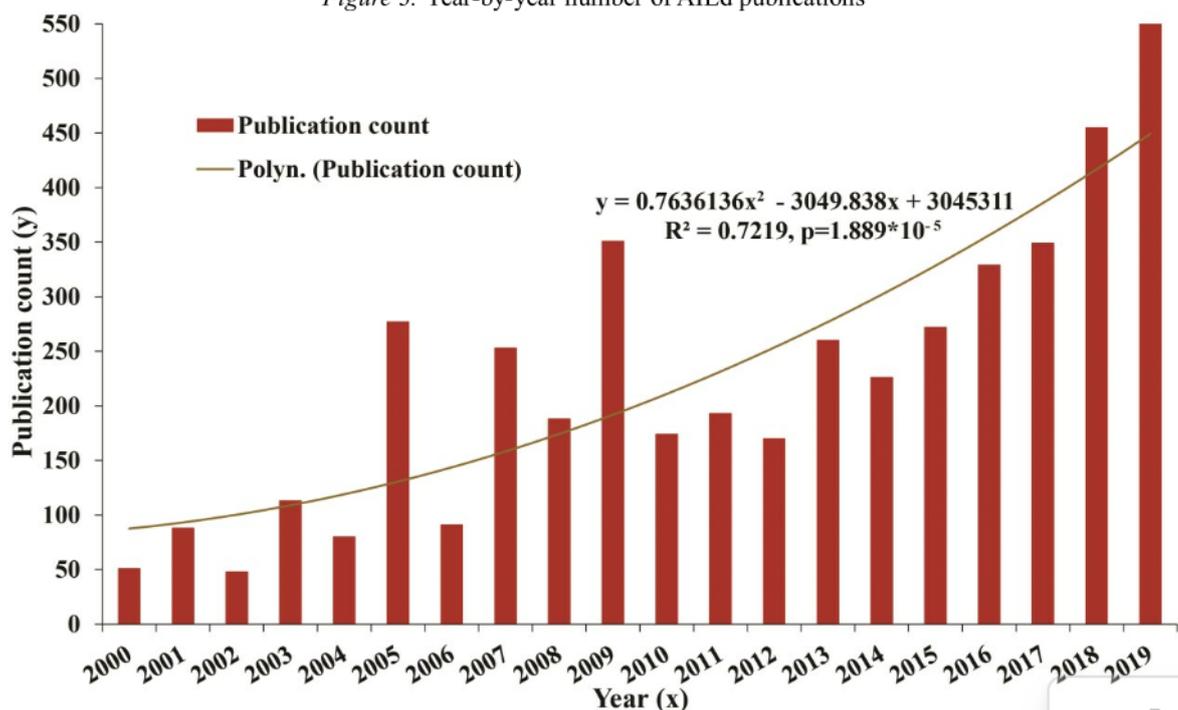
В ближайшие годы можно ожидать, что будут появляться новые законодательные инициативы, направленные на регулирование ИИ, что станет важным шагом в направлении повышения уровня социальной ответственности технологий [2]. Правительством Республики Казахстан принята Концепция по развитию искусственного интеллекта на 2024-2029 годы.

### **Международный опыт**

Опыт Китая в области этики ИИ представляет интерес. Правительство страны активно разрабатывает регуляции, направленные на контроль за использованием ИИ и защиту прав граждан. Например, в научной статье об ИИ в образовании исследователи из Китая привели анализ (Рисунок 1) о растущем количестве научных статей, опубликованных по темам «ИИ» и «Образование» от Web of Science и Google Academic с 2000 года по 2019 годы [6]. Китай также активно инвестирует в обучение специалистов в области этики ИИ, что создает основу для более ответственного использования технологий.

В Казахстане, например, обсуждаются инициативы по внедрению стандартов этического использования ИИ в рамках национальной стратегии цифровизации. Важно, чтобы страны не только перенимали зарубежный опыт, но и адаптировали его под свои реалии.

Figure 3. Year-by-year number of AIED publications



**Рисунок 1** – Анализ количества научных работ по теме ИИ в образовании по годам [6]

### Заключение

Анализ показал, что этика искусственного интеллекта требует комплексного подхода к регулированию и разработке стандартов. Баланс между инновациями и социальной ответственностью возможен только при активном сотрудничестве всех заинтересованных сторон. Прогнозируя развитие тенденций, можно отметить, что в ближайшие годы будет увеличиваться внимание к этическим аспектам ИИ, что будет способствовать его более безопасному и ответственному использованию.

В итоге, необходимо понимать, что технологический прогресс не должен ставить под угрозу социальные и этические нормы. Каждая инновация должна оцениваться с точки зрения ее влияния на общество, а разработка ИИ должна основываться на принципах этики и социальной ответственности.

### Рекомендации

1. Создание общепринятые принципы, регулирующие использование ИИ в различных сферах, включая обязательные требования к прозрачности алгоритмов.

2. Важно привлекать к обсуждению представителей разных сфер, чтобы учесть различные точки зрения и повысить уровень доверия к ИИ.

3. Повышение уровня образования по вопросам ИИ для всех слоев населения поможет снизить уровень недоверия и обеспечить более осознанное использование технологий.

4. Создание независимых органов для мониторинга и оценки использования ИИ и его воздействия на общество, что позволит выявлять и предотвращать негативные последствия.

### **Список использованных источников**

1. Миндигулова А.А. Этика и искусственный интеллект: проблемы и противоречия // Медицина. Социология. Философия. Прикладные исследования. 2022. №3. URL: <https://cyberleninka.ru/article/n/etika-i-iskusstvennyy-intellekt-problemy-i-protivorechiya>

2. Гончарова М.С. ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ: ЭТИКА И СОЦИАЛЬНАЯ ОТВЕТСТВЕННОСТЬ ЕГО ИСПОЛЬЗОВАНИЯ / М.С. Гончарова // Химия, физика, биология, математика: теоретические и прикладные исследования: сб. ст. по материалам LXXVII Международной научно-практической конференции «Химия, физика, биология, математика: теоретические и прикладные исследования». – № 10(60). – М., Изд. «Интернаука», 2023.

3. Полонская Ю. Мировое сообщество ищет способы регулировать использование технологий ИИ / Полонская Ю. [Электронный ресурс] // Казахстанская правда : [сайт]. — URL: <https://kazpravda.kz/n/mirovoe-soobshchestvo-ishchet-sposoby-regulirovat-ispolzovanie-tehnologiy-ii/>

4. Решетникова М.С., Лукина Ю.Д. Политика Китая в борьбе за мировое лидерство в области искусственного интеллекта // Вопросы инновационной экономики. 2020. №4. URL: <https://cyberleninka.ru/article/n/politika-kitaya-v-borbe-za-mirovoe-liderstvo-v-oblasti-iskusstvennogo-intellekta>

5. Мирзаев Р. Права человека в эпоху искусственного интеллекта: правовые гарантии конфиденциальности и безопасности // Universum: экономика и юриспруденция. 2024. №10 (120). URL: <https://cyberleninka.ru/article/n/prava-cheloveka-v-epohu-iskusstvennogo-intellekta-pravovye-garantii-konfidentsialnosti-i-bezopasnosti>

6. Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022). Two Decades of Artificial Intelligence in Education: Contributors, Collaborations, Research Topics, Challenges, and Future Directions. *Educational Technology & Society*, 25(1), 28–47.

7. UN E-Government Survey 2024 [Электронный ресурс]: (2024), E.24.II.H.1, United Nations, URL: <https://publicadministration.un.org/egovkb/en-us/#0>