

THE PROBLEM OF ETHICAL SUBJECTIVITY IN ARTIFICIAL INTELLIGENCE ALGORITHMS: A HUMANISTIC APPROACH

Nomuratova Makhliyokhon Kushaliyevna
University of Military Security and Defence of the Republic of Uzbekistan,
docent, Department of National Pride and Military Patriotism

Annotation. The article analyzes the problem of ethical subjectivity in artificial intelligence (AI) algorithms from the standpoint of a humanistic approach. It examines the mechanisms of algorithmic bias, the problem of decision-making transparency ("black box"), and issues of moral responsibility distribution. The author justifies the need to integrate anthropocentric principles into the development of intelligent systems to ensure their alignment with human values.

Key words: artificial intelligence, ethical subjectivity, humanism, algorithmic bias, black box, digital ethics, anthropocentrism, responsibility.

ПРОБЛЕМА ЭТИЧЕСКОЙ СУБЪЕКТИВНОСТИ В АЛГОРИТМАХ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: ГУМАНИСТИЧЕСКИЙ ПОДХОД

Номуратова Махлиёхон Кушалиевна
Университет военной безопасности и обороны Республики
Узбекистан, доцент кафедры национальной
гордости и военного патриотизма

Аннотация. В статье анализируется проблема этической субъективности в алгоритмах искусственного интеллекта (ИИ) с позиций гуманистического подхода. Рассматриваются механизмы возникновения алгоритмических предубеждений, проблема прозрачности принятия решений («черный ящик») и вопросы распределения моральной ответственности. Автором обосновывается необходимость внедрения антропоцентрических принципов в процесс разработки интеллектуальных систем для обеспечения их соответствия человеческим ценностям.

Ключевые слова: искусственный интеллект, этическая субъективность, гуманизм, алгоритмическая предвзятость, черный ящик, цифровая этика, антропоцентризм, ответственность.

В условиях стремительной цифровизации всех сфер общественной жизни искусственный интеллект (ИИ) перестал быть исключительно технологическим феноменом, превратившись в фактор, определяющий социальную и этическую реальность. Проблема «этической субъективности» алгоритмов сегодня выходит на первый план, так как системы ИИ все чаще принимают решения, затрагивающие фундаментальные права и свободы человека: от распределения социальных ресурсов до судебных решений и медицинской диагностики. Актуальность исследования обусловлена тем, что алгоритмические системы, несмотря на их кажущуюся объективность, часто наследуют скрытые предубеждения (algorithmic bias) своих создателей или данных, на которых они обучались [1].

Гуманистический подход к данной проблеме требует переосмысления роли человека в системе «человек–машина». Если классическая этика рассматривает субъекта как носителя свободной воли и моральной ответственности, то в случае с ИИ возникает феномен «распределенной ответственности», где грань между технической ошибкой и этическим выбором становится размытой. В контексте стратегии развития инноваций в Новом Узбекистане, вопросы этической безопасности ИИ и его соответствия национальным гуманистическим ценностям приобретают особую значимость для построения справедливого цифрового общества [2].

Проблема этики ИИ активно обсуждается в трудах зарубежных и отечественных философов. Теоретические основы этического программирования заложены в работах Ника Бострома и Элиезера Юдковского, которые акцентируют внимание на проблеме «выравнивания» (alignment problem) целей ИИ с человеческими ценностями [3]. Вопросы алгоритмической справедливости и социальной дискриминации подробно

анализируются в исследованиях Кэти О'Нил, которая характеризует математические модели без этического контроля как «оружие математического поражения» [4].

Среди ученых Узбекистана и стран СНГ данная проблематика начинает приобретать междисциплинарный характер, сочетая философский антропоцентризм с нормативно-правовыми аспектами цифровизации. Однако, вопрос о том, как именно гуманистические ценности могут быть операционализированы в программном коде, остается открытым и требует дальнейшего концептуального анализа.

Исследование базируется на междисциплинарном подходе, объединяющем этическую философию, когнитивистику и критический анализ алгоритмических систем. В работе использованы следующие методы:

Аксиологический анализ: для выявления системы ценностей, заложенных в современные модели ИИ, и их сопоставления с гуманистическими идеалами (справедливость, достоинство, автономия).

Дедуктивный метод: позволил перейти от общих этических принципов к частным случаям алгоритмической предвзятости в системах машинного обучения.

Критический дискурс-анализ: применен для изучения «техно-оптимистических» и «техно-пессимистических» концепций, определяющих этический статус алгоритмов в современной научной литературе [5].

Компаративный анализ: сопоставление подходов к регулированию ИИ в различных правовых и культурных системах (ЕС, Китай, США) с целью поиска универсального гуманистического стандарта. Проблема этической субъективности в ИИ проистекает из фундаментального парадокса: алгоритм математически объективен в вычислениях, но контекстуально субъективен в результатах. Проведенный анализ позволил выделить следующие ключевые аспекты:

1. Механизм трансляции человеческих предубеждений в код. Алгоритмы обучаются на данных, созданных людьми. Если исторические

данные содержат социальные, гендерные или этнические искажения, ИИ не просто копирует их, но и масштабирует. Гуманистический подход требует признания того, что ИИ не является «чистым разумом» (*tabula rasa*), а выступает зеркалом социальных противоречий. Исследование показывает, что субъективность в алгоритмах — это не техническая ошибка, а результат отсутствия этического фильтра на этапе отбора данных [6].

2. Проблема «Черного ящика» (*Black Box*) и утрата человеческого контроля. Сложные нейронные сети принимают решения, логику которых часто не могут объяснить даже их создатели. С позиции гуманизма это подрывает принцип ответственности: если решение машины непрозрачно, оно не может считаться этичным. В работе подчеркивается необходимость перехода к «объяснимому ИИ» (*XAI — Explainable AI*), где прозрачность алгоритма становится базовым моральным требованием.

3. Этическая субъективность как вопрос выбора приоритетов. Примером может служить классическая «проблема вагонетки» в беспилотных автомобилях. Алгоритм должен сделать выбор между жизнями разных категорий людей. Здесь субъективность проявляется в том, какие ценности программисты ставят на первое место: утилитаризм (минимизация общего числа жертв) или деонтологию (защита прав конкретной личности). Гуманистическая альтернатива предлагает отказ от чисто количественных оценок в пользу качественной защиты человеческого достоинства [7].

4. Пути гуманизации алгоритмов. В ходе исследования было определено, что гуманизация ИИ возможна через внедрение принципа «*Human-in-the-loop*» (Человек в цикле). Это означает, что решающее слово в критически важных для жизни человека вопросах всегда должно оставаться за субъектом, обладающим моральным сознанием и эмпатией, которыми машина лишена по своей природе.

Подводя итог исследованию проблемы этической субъективности в алгоритмических системах, можно сформулировать следующие концептуальные выводы:

Во-первых, технологическая природа искусственного интеллекта не является ценностно-нейтральной. Алгоритмы, будучи продуктом человеческой деятельности, неизбежно наследуют и масштабируют социальные предубеждения, заложенные в массивах данных. Таким образом, этическая субъективность в цифровой среде — это не временный технический сбой, а фундаментальный вызов, требующий глубокой философской рефлексии и разработки «гуманистических фильтров» на каждом этапе проектирования систем.

Во-вторых, преодоление проблемы «черного ящика» является обязательным условием для формирования доверия к ИИ. С позиции гуманизма, право человека на понимание логики принятия решений, затрагивающих его интересы, должно стать незыблемым этическим стандартом. Переход от слепого доверия к алгоритмической прозрачности позволит минимизировать риски дискриминации и обеспечит справедливое распределение ответственности в системе «человек–машина».

В-третьих, стратегическим решением проблемы остается сохранение антропоцентрического подхода. Искусственный интеллект должен выступать не как автономный субъект морали, а как когнитивный инструмент, расширяющий возможности человека при обязательном сохранении последнего слова за субъектом, обладающим эмпатией и моральной совестью.

В завершение стоит отметить, что создание этического искусственного интеллекта в Узбекистане требует интеграции глобальных технологических стандартов с национальными гуманистическими ценностями. Формирование четкой этической платформы станет гарантом того, что цифровая трансформация общества будет служить интересам личности, сохраняя её достоинство и автономия в эпоху повсеместной алгоритмизации.

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. Мирзиёев Ш.М. Стратегия «Цифровой Узбекистан – 2030». Указ Президента Республики Узбекистан.

2. Рассел С. Совместимость: Как искусственный интеллект должен отвечать целям человека. – М.: Альпина нон-фикшн, 2020. – С. 112-115.
3. Bostrom N. Superintelligence: Paths, Dangers, Strategies. – Oxford University Press, 2014. – P. 228-235.
4. O'Neil C. Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. – Broadway Books, 2016. – P. 45-52.
5. Floridi L. The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities. – Oxford University Press, 2023.
6. Floridi L. The Ethics of Artificial Intelligence: Principles, Challenges, and Opportunities. – Oxford University Press, 2023. – P. 78-82.
7. Noble S.U. Algorithms of Oppression: How Search Engines Reinforce Racism. – NYU Press, 2018. – P. 21-25.
8. Wallach W., Allen C. Moral Machines: Teaching Robots Right from Wrong. – Oxford University Press, 2008.