Дата поступления рукописи в редакцию: 16.08.2026 г.
Дата принятия рукописи в печать: 25.09.2026 г.

DOI: 10.24412/2782-3849-2026-9-67-72

ДЖАПАРОВ Ильгиз Абылкасымович,
аспирант кафедры международного права ФГАОУ ВО «Российский университет дружбы народов имени Патриса Лумумбы»,
e-mail: ilgiz_djaparov@mail.ru

Научный руководитель:
ДЕМЕНТЬЕВ А.А.,
кандидат юридических наук, доцент кафедры международного права,
e-mail: mail@law-books.ru

О ПЕРСПЕКТИВАХ ФОРМИРОВАНИЯ ЭФФЕКТИВНОЙ МЕЖДУНАРОДНО-ПРАВОВОЙ МОДЕЛИ ОБЕСПЕЧЕНИЯ ЭТИЧНОСТИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: ОТ ЭТИЧЕСКИХ ПРИНЦИПОВ К ПРАВОВЫМ НОРМАМ

Аннотация. *Статья посвящена проблеме формирования международно-правовой модели обеспечения этичности искусственного интеллекта. Автор исходит из того, что стремительное развитие технологий ИИ опережает создание юридически обязывающих норм, что порождает разрыв между этическими ожиданиями и правовой реальностью. На основе анализа ключевых международных актов - Рекомендации ЮНЕСКО об этических аспектах искусственного интеллекта 2021 года, обновленных Принципов ОЭСР по ИИ 2024 года, Рамочной конвенции Совета Европы об искусственном интеллекте 2024 года и Регламента ЕС об искусственном интеллекте - автор прослеживает эволюцию от декларативных этических принципов к формированию гибридной модели регулирования, сочетающей инструменты мягкого и жесткого права. Особое внимание уделяется принципу предосторожности и механизмам «отвердевания» мягкого права. Автор приходит к выводу, что эффективная международно-правовая модель обеспечения этичности ИИ должна строиться на основе многоуровневой архитектуры, включающей рамочные конвенции, региональные регламенты и глобальные этические стандарты, а также на институциональных механизмах мониторинга и имплементации.*

Ключевые слова: *искусственный интеллект, этика ИИ, международное публичное право, мягкое право, жесткое право, принцип предосторожности, регулирование технологий, Рамочная конвенция Совета Европы, Рекомендация ЮНЕСКО, Принципы ОЭСР.*

DZHAPAROV Ilgiz A.,
postgraduate student at the Department of International Law, Peoples' Friendship University of Russia named after Patrice Lumumba

Supervisor:
DEMENTIEV A.A.,
Candidate of Law, Associate Professor, Department of International Law

ABOUT THE PROSPECTS FOR THE FORMATION OF AN EFFECTIVE INTERNATIONAL LEGAL MODEL FOR ENSURING THE ETHICS OF ARTIFICIAL INTELLIGENCE: FROM ETHICAL PRINCIPLES TO LEGAL NORMS

Annotation. *The article addresses the problem of forming an international legal model for ensuring the ethics of artificial intelligence. The author proceeds from the fact that the rapid development of AI technologies outpaces the creation of legally binding norms, which generates a gap between ethical expectations and legal reality. Based on an analysis of key international instruments - the UNESCO Recommendation on the Ethics of Artificial Intelligence (2021), the updated OECD AI Principles (2024), the Council of Europe Framework Convention on Artificial Intelligence (2024), and the EU Artificial Intelligence Act - the author traces the evolution from declarative ethical principles to the formation of a hybrid regulatory model combining soft law and*

hard law instruments. Particular attention is paid to the precautionary principle and the mechanisms of "hardening" of soft law. The author concludes that an effective international legal model for ensuring AI ethics should be built on a multi-level architecture, including framework conventions, regional regulations, and global ethical standards, as well as institutional mechanisms for monitoring and implementation.

Key words: *artificial intelligence, AI ethics, public international law, soft law, hard law, precautionary principle, technology regulation, Council of Europe Framework Convention, UNESCO Recommendation, OECD Principles.*

Введение

Одной из наиболее острых проблем, с которыми сталкивается современное международное публичное право, является необходимость регулирования технологий, этические импликации которых значительно опережают возможности правового реагирования. Искусственный интеллект, ставший за последние два десятилетия неотъемлемой частью экономических, социальных и политических процессов, порождает фундаментальный вопрос: каким образом международное право может обеспечить этичность его разработки и применения, не препятствуя при этом технологическому прогрессу и не пытаясь дать окончательные ответы на вопросы, которые остаются предметом незавершённых научных дискуссий?

Сложность ситуации придаёт тот факт, что само понятие искусственного интеллекта не обрело общепризнанного научного наполнения. Инженерные дисциплины, когнитивная наука и философия сознания предлагают различные трактовки интеллектуального поведения - от оперирования абстрактными символами по строгим логическим правилам до незапрограммированных свойств, рождающихся внутри многослойных искусственных нейронов. Эта многозначность создаёт ситуацию, при которой объект регулирования ускользает от строгого определения именно в силу своей научной неопределённости. Как отмечается в доктрине, отсутствие единого понимания сущности искусственного интеллекта ставит перед правом задачу, которую невозможно решить традиционными средствами: регулировать необходимо, а опереться на готовое знание невозможно [1].

Вместе с тем этическая проблематика искусственного интеллекта не сводится к терминологическим спорам. Речь идёт о вполне конкретных рисках: алгоритмической дискриминации, непрозрачности решений, используемых в сферах здравоохранения, кредитования, правосудия и трудоустройства, манипуляции общественным мнением, распространении дезинформации, а также о более отдалённой перспективе утраты человеческого контроля над критически важными системами [2]. Именно эти риски, а не абстрактные философские разногласия, определяют повестку международного правотворчества.

Таким образом, международное право оказалось перед необходимостью выработать регуляторную стратегию, которая, с одной стороны, признавала бы границы доступного сегодня знания, а с другой - обеспечивала бы достаточный уровень защиты этических ценностей и прав человека. Настоящая статья ставит перед собой цель проанализировать формирующуюся международно-правовую модель обеспечения этичности искусственного интеллекта и оценить перспективы её дальнейшего развития.

Цель и задачи

Цель настоящей статьи - показать, что международно-правовое регулирование этичности искусственного интеллекта развивается в направлении гибридной модели, сочетающей инструменты мягкого и жёсткого права, и обосновать, что такая модель является не временным компромиссом, а закономерной и наиболее адаптивной формой регулирования в условиях научной и технологической неопределённости.

Для достижения поставленной цели необходимо последовательно решить три задачи. Во-первых, проследить эволюцию международных подходов к этическому регулированию ИИ - от первых декларативных принципов к формированию юридически обязывающих инструментов. Во-вторых, проанализировать ключевые международные акты - Рекомендацию ЮНЕСКО 2021 года, обновлённые Принципы ОЭСР 2024 года, Рамочную конвенцию Совета Европы 2024 года и Регламент ЕС об искусственном интеллекте - с точки зрения их вклада в формирование эффективной модели обеспечения этичности. В-третьих, обосновать, что дальнейшее развитие международно-правовой модели должно опираться на многоуровневую архитектуру, включающую рамочные конвенции, региональные регламенты, глобальные этические стандарты и институциональные механизмы мониторинга.

Методология

Методологический инструментарий настоящей статьи носит комплексный характер, что обусловлено междисциплинарным характером исследуемой проблематики, лежащей на стыке международного публичного права, философии права и исследований в области регулирования технологий. Центральное место занимает системно-структурный подход, позволивший рассмот-

реть формирующуюся международно-правовую модель обеспечения этичности ИИ не как совокупность разрозненных актов, а как целостную систему, элементы которой - мягкое право, жёсткое право и институциональные механизмы - находятся в отношениях взаимодополняемости.

Сравнительно-правовой метод применён при сопоставлении подходов к этическому регулированию ИИ, отражённых в документах различной юридической силы и географического охвата - от глобальных рекомендаций ЮНЕСКО до региональных регламентов ЕС. Формально-юридический метод использован при анализе текстов международных актов и определении их юридической природы. Элементы метода правового моделирования задействованы при обосновании перспектив формирования многоуровневой международно-правовой модели. Наконец, историко-правовой метод позволил проследить эволюцию от первых этических деклараций к юридически обязывающим инструментам.

Результаты

Первая волна международных инициатив в сфере этики искусственного интеллекта, приходящаяся на 2017–2021 годы, характеризовалась преобладанием декларативных документов, закреплявших общие ценностные ориентиры без создания механизмов их имплементации. Как отмечается в аналитических обзорах, этот период породил «пролиферацию принципов, но не институционализацию» [3]. Асиломарские принципы ИИ 2017 года, руководящие принципы G20 2019 года, Этические руководящие принципы для надёжного ИИ, разработанные экспертной группой высокого уровня Европейской комиссии в 2019 году, - все эти документы, при всей их значимости для формирования международного консенсуса относительно ценностей, не создавали юридических обязательств и не предусматривали механизмов мониторинга [4, с.49].

Ситуация принципиально изменилась в 2021 году с принятием Рекомендации ЮНЕСКО об этических аспектах искусственного интеллекта - первого глобального стандарта в этой сфере, одобренного всеми 194 государствами-членами организации. Рекомендация закрепила фундаментальные принципы - защиту прав человека и достоинства, прозрачность, справедливость, человеческий контроль над системами ИИ - и, что особенно важно, предложила развёрнутые направления политических действий, позволяющие государствам переводить эти принципы в конкретные меры в области управления данными, окружающей среды, гендерного равенства, образования и здравоохранения. Несмотря на то, что Рекомендация носит рекомендательный характер, её значение для формирования международно-правовой модели обеспечения этичности ИИ трудно переоценить: она создала нормативную рамку, на которую впоследствии стали опираться как региональные, так и национальные регуляторы.

Параллельно развивалась линия ОЭСР. Принципы ОЭСР по искусственному интеллекту, впервые принятые в 2019 году [5], были существенно обновлены в мае 2024 года [6]. Обновлённая версия, сохраняя приверженность ценностям инновационного и надёжного ИИ, уважающего права человека и демократические ценности, более прямо адресовала вызовы, связанные с приватностью, интеллектуальной собственностью, безопасностью и целостностью информации. Принципы ОЭСР, насчитывающие 47 приверженцев, включая ЕС, выполняют функцию глобального ориентира для политических рамок в сфере ИИ. Особенно важно, что определения системы ИИ и её жизненного цикла, закреплённые в Рекомендации ОЭСР, используются в качестве основы для нормативных актов Европейского союза, Японии и Соединённых Штатов, а также в работе Организации Объединённых Наций [7]. Тем самым мягкое право ОЭСР выполняет функцию нормативной гармонизации, создавая общий понятийный аппарат для последующего жёсткого регулирования.

Качественный сдвиг в формировании международно-правовой модели обеспечения этичности ИИ произошёл в 2024 году с принятием Рамочной конвенции Совета Европы об искусственном интеллекте и правах человека, демократии и верховенстве права (CETS № 225) [8]. Конвенция, открытая для подписания 5 сентября 2024 года, стала первым в мире юридически обязывающим международным договором, направленным на обеспечение соответствия деятельности в жизненном цикле систем ИИ общим стандартам в области прав человека, демократии и верховенства права.

Конвенция восприняла риск-ориентированный подход, предполагающий оценку серьёзности и вероятности негативного воздействия систем ИИ на права человека, демократию и верховенство права. Для этого разрабатывается специальная методология HUDERIA (Human Rights, Democracy and Rule of Law Impact Assessment), призванная обеспечить систематическую идентификацию, оценку и минимизацию рисков. Принципиально важно, что Конвенция не пытается дать содержательное определение искусственного интеллекта, а оперирует функциональными характеристиками систем ИИ, что позволяет избежать ловушки терминологических споров, о которой говорилось выше.

Вместе с тем, как отмечается в критической литературе, Рамочная конвенция остаётся региональным инструментом и не охватывает в

полной мере экологические аспекты регулирования ИИ. Кроме того, её эффективность будет зависеть от имплементации на национальном уровне, что создаёт риск фрагментации регулирования.

Регламент (ЕС) 2024/1689, вступивший в силу 2 августа 2024 года, представляет собой наиболее разработанный и всеобъемлющий пример жёсткого правового регулирования в сфере ИИ [9]. Регламент ставит целью обеспечение разработки и внедрения безопасных и надёжных систем ИИ на едином рынке ЕС при одновременном обеспечении здоровья и безопасности граждан и уважения основных прав.

Ключевой особенностью Регламента является риск-ориентированный подход: чем выше риск причинения вреда обществу, тем строже правила. Регламент определяет ряд практик ИИ как недопустимых, включая сублиминальные или обманные техники манипуляции поведением, эксплуатацию уязвимостей, обусловленных возрастом, инвалидностью или социально-экономическим положением, а также социальный скоринг. Высокорисковые системы ИИ, используемые в критической инфраструктуре, образовании, трудоустройстве, основных услугах, правоприменении, миграционном контроле и отправлении правосудия, подлежат строгим требованиям и обязательной оценке соответствия [10].

Этическая составляющая Регламента проявляется через закрепление семи принципов надёжного ИИ: человеческое агентство и надзор, техническая надёжность и безопасность, приватность и управление данными, прозрачность, разнообразие, недискриминация и справедливость, социальное и экологическое благополучие, подотчётность. Таким образом, Регламент ЕС представляет собой попытку перевести этические принципы, ранее закреплённые в мягко-правовых документах, в юридически обязывающие нормы с конкретными механизмами имплементации и санкциями за нарушение.

Анализ рассмотренных актов позволяет сделать вывод, что формирующаяся международно-правовая модель обеспечения этичности ИИ носит гибридный характер, сочетая инструменты мягкого и жёсткого права. Мягкое право - Рекомендация ЮНЕСКО, Принципы ОЭСР, руководящие принципы G7 - выполняет несколько критически важных функций. Оно обеспечивает гибкость и адаптивность регулирования в условиях стремительного технологического развития, позволяет достичь широкого международного консенсуса относительно ценностных ориентиров, а также создаёт понятийный аппарат и нормативную рамку для последующего жёсткого регулирования.

Вместе с тем мягкое право имеет очевидные ограничения. Как отмечается в исследованиях, необязательные меры ограничены в том, что могут достичь: в отсутствие механизмов принуждения они не способны предотвратить или устранить вред, связанный с ИИ, в критически важных сферах. Именно поэтому наблюдается тенденция к «отвердеванию» мягкого права - постепенному превращению рекомендательных норм в элементы жёсткого регулирования по мере накопления практики и формирования консенсуса. Рамочная конвенция Совета Европы и Регламент ЕС об ИИ являются примерами такого «отвердевания», поскольку они переводят принципы, ранее закреплённые в мягко-правовых документах, в юридически обязывающие нормы.

Методологическим фундаментом такого подхода выступает принцип предосторожности, реципированный международным правом из экологической сферы. В условиях, когда научное понимание рисков, связанных с ИИ, остаётся неполным, а потенциальный ущерб может быть серьёзным и необратимым, упреждающие запреты на определённые практики - такие как социальный скоринг или отдельные методы биометрической идентификации - представляют собой не чрезмерную осторожность, а осознанную стратегию управления неизвестностью. Как показывает опыт Регламента ЕС, логика предосторожности позволяет вводить запреты на потенциально вредоносные практики ещё до того, как их вредоносность будет исчерпывающе доказана.

Проведённый анализ позволяет сформулировать ряд выводов относительно перспектив формирования эффективной международно-правовой модели обеспечения этичности искусственного интеллекта.

Во-первых, такая модель должна носить многоуровневый характер. Глобальный уровень представлен этическими стандартами ЮНЕСКО и принципами ОЭСР, обеспечивающими универсальную ценностную рамку и понятийный аппарат. Региональный уровень - Рамочная конвенция Совета Европы, Регламент ЕС - обеспечивает юридически обязывающие механизмы с конкретными процедурами оценки рисков и имплементации. Национальный уровень, в свою очередь, адаптирует эти нормы к конкретным правовым системам и практикам.

Во-вторых, эффективная модель должна сочетать гибкость мягкого права с обязательностью жёсткого. Мягкое право обеспечивает адаптивность регулирования в условиях технологической неопределённости, тогда как жёсткое право создаёт механизмы принуждения и ответственности. Динамическое взаимодействие между этими уровнями - «отвердевание» мягкого права по мере формирования консенсуса и накопления практики - является ключевым механизмом эволюции

регулирования [11].

В-третьих, необходимо развитие институциональных механизмов мониторинга и имплементации. Создание Независимой международной научной группы по ИИ, обсуждаемое в рамках ООН [12], а также механизмов глобального диалога по управлению ИИ, может обеспечить постоянный мониторинг технологического развития и выработку научно обоснованных рекомендаций для политиков. Такие механизмы должны быть независимыми от политического влияния, включать экспертов из широкого спектра дисциплин и обеспечивать сбалансированное географическое и гендерное представительство.

В-четвёртых, особое внимание должно уделяться проблеме «этического отмывания» (ethics-washing) - использованию этической риторики для создания видимости ответственного поведения без реальных изменений в практике. Эффективная модель должна предусматривать механизмы верификации и аудита, позволяющие отличать подлинную приверженность этическим принципам от их имитации.

Заключение

Формирование международно-правовой модели обеспечения этичности искусственного интеллекта представляет собой сложный и многоаспектный процесс, отражающий фундаментальное противоречие между необходимостью регулирования и невозможностью опереться на окончательное научное знание о сущности регулируемого явления. Проведённый анализ показывает, что международное право отвечает на эту ситуацию не отказом от регулирования, а выработкой гибридных стратегий, сочетающих мягкое и жёсткое право, рамочные конвенции и региональные регламенты, ценностные ориентиры и процедурные механизмы.

Ключевые элементы формирующейся модели - Рекомендация ЮНЕСКО как глобальный этический стандарт, Принципы ОЭСР как инструмент нормативной гармонизации, Рамочная конвенция Совета Европы как первый юридически обязывающий международный договор в сфере ИИ и Регламент ЕС как наиболее разработанная система жёсткого регулирования - образуют многоуровневую архитектуру, способную адаптироваться к технологическим изменениям, не отказываясь при этом от защиты фундаментальных этических ценностей.

Перспективы дальнейшего развития этой модели связаны с укреплением институциональных механизмов мониторинга и имплементации, развитием методологий оценки рисков, обеспечением подлинной, а не декларативной приверженности этическим принципам, а также с расширением географического охвата юридически обязывающих инструментов за пределы европейского региона. В условиях, когда технологическое развитие продолжает опережать правовое регулирование, именно такая - гибкая, многоуровневая и адаптивная - модель представляется наиболее реалистичным и эффективным ответом на вызовы, порождаемые искусственным интеллектом.

Список литературы:

[1] Легитимация этических стандартов правового регулирования искусственного интеллекта в практике отечественного правотворчества и правоприменения // КиберЛенинка. URL: https://cyberleninka.ru/article/n/legitimatsiya-eticheskih-standartov-pravovogo-regulirovaniya-iskusstvennogo-intellekta (дата обращения: 15.07.2026).

[2] Рекомендация об этических аспектах искусственного интеллекта: принята 23 нояб. 2021 г. Генеральной конференцией ЮНЕСКО на 41-й сессии. Париж : ЮНЕСКО, 2021. URL: https://unesdoc.unesco.org/ark:/48223/pf0000377897 (дата обращения: 15.07.2026).

[3] Bridging Soft and Hard Law in AI Governance. BIICL, 2025. URL: https://www.biicl.org/blog/121/upcoming-courses (дата обращения: 24.08.2026).

[4] Лисаченко А.В., Берг Л.Н. Правовое обеспечение безопасного использования генетической и геномной информации: учебник для вузов, Москва: Издательство Юрайт, 2026. С. 48–50.

[5] Рекомендация Совета ОЭСР по искусственному интеллекту // Юридические инструменты ОЭСР. URL: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449 (дата обращения: 23.08.2026).

[6] OECD updates AI Principles to stay abreast of rapid technological developments. OECD, 03.05.2024. URL: https://www.oecd.org/en/about/news/press-releases/2024/05/oecd-updates-ai-principles-to-stay-abreast-of-rapid-technological-developments.html (дата обращения: 26.08.2026).

[7] Enablers, guardrails and engagement for unlocking trustworthy AI: Governing with Artificial Intelligence. OECD, 2025. URL: https://www.oecd.org/en/publications/governing-with-artificial-intelligence_795de142-en.html (дата обращения: 01.09.2026).

[8] Рамочная конвенция Совета Европы об искусственном интеллекте и правах человека, демократии и верховенстве права (CETS № 225) // Официальный сайт Совета Европы. URL: https://rm.coe.int/1680afae3c (дата обращения: 01.09.2026).

[9] Регламент (ЕС) 2024/1689 Европейского парламента и Совета от 13 июня 2024 года, устанавливающий гармонизированные правила в области искусственного интеллекта // EUR-Lex. URL: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/

eng (дата обращения: 03.09.2026).

[10] Кашкин С. Ю. Право Европейского союза : учебник для вузов. 4-е изд., перераб. и доп. Москва : Издательство Юрайт, 2026. С. 54–55.

[11] Hybrid global governance for responsible and inclusive Artificial Intelligence: Proposing a new Sustainable Development Goal 18 // ScienceDirect. 2025. URL: https://www.sciencedirect.com/science/article/pii/S0160791X25003495 (дата обращения: 05.09.2026).

[12] EU Statement – UN General Assembly: Informal Consultations on Artificial Intelligence. EEAS, 17.01.2025. URL: https://www.eeas.europa.eu/delegations/un-new-york/eu-statement-artificial-intelligence_en (дата обращения: 05.09.2026).

References:

[1] Legitimization of ethical standards for the legal regulation of artificial intelligence in the practice of domestic law-making and law enforcement // CyberLeninka. URL: https://cyberleninka.ru/article/n/legitimatsiya-eticheskih-standartov-pravovogo-regulirovaniya-iskusstvennogo-intellekta (дата обращения: 15.07.2026).

[2] Recommendation on Ethical Aspects of Artificial Intelligence: Adopted Nov. 23 2021 by the UNESCO General Conference at the 41st session. Paris: UNESCO, 2021. URL: https://unesdoc.unesco.org/ark:/48223/pf0000377897 (accessed on: 15.07.2026).

[3] Bridging Soft and Hard Law in AI Governance. BIICL, 2025. URL: https://www.biicl.org/blog/121/upcoming-courses (access date: 24.08.2026).

[4] Lisachenko A.V., Berg L.N. Legal support for the safe use of genetic and genomic information: a textbook for universities, Moscow: Yurayt Publishing House, 2026. S. 48-50.

[5] OECD Council Recommendation on Artificial Intelligence // OECD Legal Instruments. URL: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449 (дата обращения: 23.08.2026).

[6] OECD updates AI Principles to stay abreast of rapid technological developments. OECD, 03.05.2024. URL: https://www.oecd.org/en/about/news/press-releases/2024/05/oecd-updates-ai-principles-to-stay-abreast-of-rapid-technological-developments.html (дата обращения: 26.08.2026).

[7] Enablers, guardrails and engagement for unlocking trustworthy AI: Governing with Artificial Intelligence. OECD, 2025. URL: https://www.oecd.org/en/publications/governing-with-artificial-intelligence_795de142-en.html (дата обращения: 01.09.2026).

[8] Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law (CETS No. 225) // Official Website of the Council of Europe. URL: https://rm.coe.int/1680afae3c (access date: 01.09.2026).

[9] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules in the field of artificial intelligence // EUR-Lex. URL: https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng (access date: 03.09.2026).

[10] Kashkin S. Yu. Law of the European Union: a textbook for universities. 4th ed., Revised and add. Moscow: Yuright Publishing House, 2026. S. 54-55.

[11] Hybrid global governance for responsible and inclusive Artificial Intelligence: Proposing a new Sustainable Development Goal 18 // ScienceDirect. 2025. URL: https://www.sciencedirect.com/science/article/pii/S0160791X25003495 (access date: 05.09.2026).

[12] EU Statement – UN General Assembly: Informal Consultations on Artificial Intelligence. EEAS, 17.01.2025. URL: https://www.eeas.europa.eu/delegations/un-new-york/eu-statement-artificial-intelligence_en (дата обращения: 05.09.2026).