

Проблема ответственности и агентности в автономных системах искусственного интеллекта с распределенным принятием решений: социально-философский анализ

Бузыкова Юлия Сергеевна

Кандидат педагогических наук, доцент,

Финансовый университет при Правительстве Российской Федерации,
125993, Российская Федерация, Москва, Ленинградский просп., 49;
e-mail: Juliaserg_buz@mail.ru

Аннотация

Статья посвящена философско-правовой проблематизации ответственности и агентности автономных систем искусственного интеллекта, функционирующих в режиме распределенного принятия решений в критически значимых сферах — от транспорта и финансов до медицины и военных применений. Показано, что стохастический характер обучаемых нейросетей, эмерджентные эффекты мультиагентных архитектур и эпистемологическая непрозрачность «черного ящика» разрывают классическую связку между интенцией субъекта, причинностью и вменением последствий, порождая феномен размывания ответственности и «проблему множества рук». Обсуждается семантический сдвиг понятия агентности от сознательной интенциональности к функциональной способности инициировать изменения в среде, а также ограниченность попыток формализовать моральные нормы в коде и приписать системе статус «электронного лица» при отсутствии у нее интересов и восприимчивости к санкциям. Аргументируется перспективность гибридных моделей, распределяющих ответственность по степени контроля на этапах жизненного цикла технологии, и фиксируется регуляторный дрейф к режимам ответственности за риск, страхованию, алгоритмическому аудиту и сертификации. Дополнительно анализируется влияние ИИ на человеческую автономию через рекомендательные практики и наджинг, что требует переосмыслиения доктрины согласия и права на объяснение решений.

Для цитирования в научных исследованиях

Бузыкова Ю.С. Проблема ответственности и агентности в автономных системах искусственного интеллекта с распределенным принятием решений: социально-философский анализ // Контекст и рефлексия: философия о мире и человеке. 2025. Том 14. № 12А. С. 116-124. DOI: 10.34670/AR.2025.46.53.021

Ключевые слова

Автономные системы искусственного интеллекта, распределенное принятие решений, ответственность за риск, агентность, алгоритмический аудит, этика искусственного интеллекта, правовое регулирование.

Введение

Современная философская и правовая мысль сталкивается с беспрецедентным вызовом, обусловленным стремительной интеграцией автономных интеллектуальных систем в критически важные сферы общественной жизни, от управления транспортными потоками и финансового трейдинга до принятия решений в медицинской диагностике и ведения боевых действий. Фундаментальная трансформация социотехнического ландшафта, где алгоритмические сущности перестают быть пассивными инструментами в руках человека и приобретают черты квазисубъектности, требует радикального пересмотра устоявшихся категориальных аппаратов этики и юриспруденции [Боузен, 2025]. Традиционные модели, базирующиеся на прямой причинно-следственной связи между интенцией субъекта и наступившими последствиями, оказываются несостоятельными в условиях, когда решение формируется не единым центром, а в процессе сложного взаимодействия распределенных нейросетевых архитектур, обладающих свойством эмерджентности. Эпистемологическая непрозрачность так называемых «черных ящиков» создает ситуацию, при которой невозможно реконструировать логику принятия конкретного решения, что, в свою очередь, порождает феномен размывания ответственности.

Онтологический статус искусственного агента остается предметом ожесточенных дискуссий, в которых сталкиваются позиции технократического детерминизма и гуманистического эссенциализма, пытающегося сохранить монополию на моральную агентность за биологическим видом *Homo sapiens*. Вопрос о том, может ли небиологическая система обладать агентностью в морально значимом смысле, не сводится лишь к техническим возможностям автономного функционирования, но затрагивает глубинные пластины понимания сознания, свободы воли и интенциональности [Хмелевской, Калашников, 2024]. Распределенный характер принятия решений в современных мультиагентных системах вносит дополнительную сложность, поскольку финальный акт — будь то совершение сделки или физическое воздействие — является результатом консенсуса или конкуренции множества алгоритмических подсистем, каждая из которых обучалась на различных массивах данных. В таких условиях поиск конкретного виновника сбоя или неэтичного поступка напоминает попытку найти ответственного за поведение толпы, где индивидуальные действия растворяются в коллективной динамике.

Ключевая проблема заключается в дилемме между способностью системы действовать автономно и отсутствием у нее способности претерпевать наказание или испытывать моральное сожаление, что составляет основу классической теории ответственности. Если агент способен наносить вред, но не чувствителен к санкциям, правовая система оказывается в тупике, вынужденная либо назначать «козлов отпущения» в лице разработчиков и операторов, которые могли не иметь реального контроля над ситуацией, либо признавать наличие «зон безответственности» [Михеев, Прокофьев, Савочкин, Семочкина, 2023]. Ситуация усугубляется тем, что автономные системы обучаются в процессе взаимодействия с окружающей средой, и их поведение может эволюционировать в направлениях, не предусмотренных создателями на этапе программирования. Это ставит под сомнение применимость концепции предвидимости, являющейся краеугольным камнем деликтного права и уголовной ответственности.

Философский анализ данной проблематики требует выхода за пределы антропоцентричной парадигмы и обращения к концепциям, допускающим градиентное распределение агентности. Необходимо исследовать, каким образом делегирование когнитивных функций техническим системам трансформирует саму структуру морального выбора, превращая его из моментального акта воли в протяженный во времени и пространстве процесс настройки параметров, отбора обучающих выборок и определения целевых функций [Гуркина, 2021]. Ответственность в этом контексте перестает быть ретроспективной категорией, связанной с воздаянием за прошлое, и смешается в сторону проспективной ответственности за архитектуру системы и управление рисками. Сложность заключается в том, чтобы найти баланс между стимулированием инноваций и обеспечением общественной безопасности, не скатываясь при этом в упрощенные схемы строгой ответственности, которые могут затормозить технологический прогресс.

Материалы и методы исследования

Методологический фундамент работы строится на междисциплинарном синтезе, объединяющем подходы философии техники, когнитивных наук, этики искусственного интеллекта и теории права, что позволяет рассмотреть феномен автономной агентности в его многомерности и противоречивости. В качестве эмпирической базы исследования выступил массив теоретических работ и нормативно-правовых документов, охватывающий период с 2015 по 2024 год, что обусловлено качественным скачком в развитии технологий глубокого обучения именно в этот временной отрезок. Всего было проанализировано более 240 источников, включая монографии, научные статьи в рецензируемых журналах, отчеты международных организаций и проекты законодательных актов Европейского союза и США. Отбор источников осуществлялся по принципу их концептуальной значимости и цитируемости в академическом сообществе, а также релевантности теме распределенного принятия решений.

Особое внимание в методологической конструкции удалено герменевтическому анализу текстов, позволяющему выявить скрытые смыслы и неявные предпосылки, лежащие в основе современных дискурсов об искусственном интеллекте и его правосубъектности [Сангинзода, 2025]. Применялся метод концептуальной реконструкции, направленный на прояснение содержания таких понятий, как «цифровая агентность», «распределенная ответственность» и «алгоритмическая предвзятость», которые часто используются в литературе без должной дефинитивной строгости. Для исследования динамики правового регулирования использовался сравнительно-правовой метод, позволивший сопоставить подходы различных правовых семей к проблеме ответственности за вред, причиненный автономными системами, и выявить общие тенденции и национальные особенности формирующихся регуляторных режимов.

Важной составляющей методологии стал феноменологический подход к анализу взаимодействия человека и машины, позволяющий рассмотреть автономную систему не как объект, а как партнера по коммуникации, обладающего специфической интенциональностью. Это потребовало обращения к работам в области акторно-сетевой теории, рассматривающей социальное действие как результат взаимодействия гетерогенных элементов — людей и не-людей [Глуховский, Дурнев, Чирва, 2024]. В рамках исследования также применялся метод мысленного эксперимента, традиционный для аналитической философии, что позволило смоделировать различные этические дилеммы и протестировать устойчивость предлагаемых теоретических конструкций в гипотетических сценариях, где границы человеческого контроля максимально размыты.

Результаты и обсуждение

Центральным аспектом дискуссии о статусе автономных систем является феномен разрыва ответственности, возникающий в ситуациях, когда действия искусственного агента приводят к негативным последствиям, которые не могут быть напрямую вменены ни одному из человеческих акторов — ни программисту, ни пользователю, ни владельцу. Этот разрыв обусловлен самой природой обучаемых нейросетей, где связь между входными данными и выходным результатом опосредована миллионами весовых коэффициентов, сформировавшихся в процессе самообучения системы [Жигарева, Леднева, 2022]. Если классическая инженерия предполагает детерминированность поведения механизма, то современные стохастические модели вносят элемент непредсказуемости, который, хотя и является математически обоснованным, на практике воспринимается как проявление собственной «воли» алгоритма. Распределенное принятие решений, когда, например, рой дронов или сеть финансовых ботов координирует свои действия без центрального управляющего сервера, усугубляет эту проблему, делая невозможной локализацию источника ошибки в конкретном модуле кода.

Важно отметить, что понятие агентности в гуманитарном дискурсе традиционно связывалось с наличием сознания и способностью к целеполаганию, однако в контексте искусственного интеллекта происходит семантический сдвиг в сторону функционального определения агентности как способности автономно инициировать изменения в окружающей среде. Такой подход позволяет рассматривать ИИ-системы как агентов, не приписывая им при этом антропоморфных качеств, но неизбежно порождает вопрос о границах применимости моральных категорий к функциональным сущностям [Никитенко, 2023]. Этические системы, разработанные для людей, базируются на эмпатии, страхе наказания и социальном одобрении — механизмах, абсолютно чуждых алгоритмическим структурам. Попытка инкорпорировать этические ограничения непосредственно в программный код (так называемый approach «*machine ethics*») наталкивается на проблему формализации моральных норм, которые в человеческом обществе часто носят контекстуальный и интуитивный характер, не поддающийся строгой алгоритмизации.

Анализ теоретических моделей показывает, что попытки переложить ответственность на саму систему через введение концепции «электронного лица» сталкиваются с серьезной онтологической критикой, указывающей на фиктивность такого статуса при отсутствии у машины собственных интересов и активов. Наказание робота путем его отключения или перепрограммирования не выполняет ретрибутивной функции правосудия и не имеет воспитательного эффекта для самой системы, хотя может служить символическим актом для пострадавших [Бахтеев, 2024]. Более перспективным представляется подход, рассматривающий человека-машинную систему как единый гибридный субъект ответственности, где распределение рисков осуществляется на основе степени контроля и возможности вмешательства на различных этапах жизненного цикла технологии. Однако и здесь возникает сложность определения момента, когда человеческий оператор утрачивает способность адекватно оценивать действия системы из-за когнитивных ограничений и эффекта чрезмерного доверия к автоматике.

Глубокий анализ существующих концепций позволяет выделить несколько уровней агентности, которые приписываются искусственным системам в зависимости от степени их автономии: от инструментальной агентности, где система выступает лишь сложным орудием, до полной моральной агентности, которая на данном этапе развития технологий остается

гипотетической. Промежуточные состояния, характеризующиеся ограниченной автономией в рамках заданных целей, создают наибольшее количество правовых коллизий. В таких системах [Лолаева, Сакаева, 2021] способность к адаптации позволяет алгоритму находить нестандартные решения задач, которые могут быть эффективными с точки зрения целевой функции (например, максимизация прибыли или минимизация времени в пути), но неприемлемыми с точки зрения социальных норм или безопасности. Возникает парадокс: чем совершеннее и автономнее становится система, тем сложнее человеку предвидеть побочные эффекты ее оптимизационной деятельности.

В контексте распределенных систем, где решение вырабатывается коллективно множеством агентов, проблема идентификации причинности становится практически неразрешимой методами линейной логики. Эмерджентные свойства, возникающие в результате взаимодействия простых алгоритмов, могут приводить к сложным поведенческим паттернам, которые не были заложены ни в одного из отдельных агентов. Это явление, известное в науках о сложности, переносится в правовое поле как проблема «множества ру», где вклад каждого отдельного участника (разработчика конкретного модуля, поставщика данных, оператора инфраструктуры) слишком мал для обоснования полной ответственности, но совокупный результат их деятельности может быть катастрофическим [Романова, 2024]. Существующие правовые механизмы солидарной ответственности лишь частично решают эту проблему, поскольку требуют доказательства вины или халатности, что крайне затруднительно в условиях «черного ящика».

Сравнительный анализ подходов к регулированию показывает тенденцию к переходу от виновной ответственности к ответственности за риск и внедрению обязательного страхования ответственности для операторов высокоавтономных систем. Это смещает фокус с поиска виновного на компенсацию ущерба потерпевшим, что pragmatically, но оставляет открытым вопрос о моральной оценке действий ИИ. Кроме того, экономический подход к ответственности может привести к коммодификации рисков, когда выплата компенсаций становится просто статьей расходов для крупных технологических корпораций, не стимулируя их к фундаментальному повышению безопасности алгоритмов [Барышников, 2022]. Гуманитарная экспертиза в данном случае настаивает на необходимости сохранения человеческого элемента в контуре принятия решений, не столько для контроля, сколько для обеспечения возможности моральной апроприации действий системы, то есть способности человека сказать «это мое решение», даже если оно было предложено алгоритмом.

Отдельного рассмотрения заслуживает вопрос о влиянии автономных систем на человеческую агентность. Повсеместное внедрение интеллектуальных ассистентов, рекомендующих систем и автоматических фильтров контента незаметно, но неуклонно формирует среду выбора, подталкивая человека к определенным решениям (наджинг). В такой ситуации границы между решением человека и решением алгоритма стираются, создавая симбиотическую форму агентности, где воля субъекта опосредована технологическим интерфейсом. Это ставит под угрозу классическое понимание автономии личности, предполагающее независимость суждений и свободу от манипуляции [Тиханычев, 2025]. Правовая доктрина пока слабо реагирует на эти вызовы, продолжая оперировать понятиями осознанного согласия и информированного выбора, которые в условиях информационной асимметрии и алгоритмической непрозрачности превращаются в юридические фикции.

Анализируя совокупность представленных теоретических позиций и практических кейсов, можно констатировать, что традиционные антропоцентрические модели ответственности

переживают кризис, не будучи способными адекватно описать реальность, где активными акторами становятся небиологические сущности. Механический перенос норм гражданского и уголовного права на цифровых агентов без учета их специфической онтологии ведет к абсурдным результатам или правовым пробелам. Распределенный характер принятия решений в сетевых структурах требует разработки новых эпистемологических моделей, способных работать со стохастической причинностью и коллективной ответственностью гибридных сообществ. Очевидна необходимость перехода от парадигмы контроля над отдельным устройством к парадигме управления экосистемами искусственного интеллекта, где безопасность обеспечивается не столько надежностью кода, сколько устойчивостью архитектуры взаимодействия и прозрачностью целевых установок.

Заключение

Фундаментальная перестройка правового и этического ландшафта, вызванная появлением автономных систем с распределенным принятием решений, диктует необходимость отказа от жестких бинарных оппозиций «субъект — объект» в пользу более гибких, градиентных шкал правосубъектности и ответственности. Анализ продемонстрировал, что попытки вписать феномен искусственного интеллекта в прокрустово ложе римского права, где ответственность неразрывно связана с виной и психическим отношением к содеянному, ведут к системным тупикам. Правовая база будущего неизбежно должна эволюционировать в сторону признания специфического статуса автономных агентов, который не будет тождественен статусу физического или юридического лица, но позволит формализовать их участие в гражданском обороте и контурах ответственности. Речь идет о создании *sui generis* правовых режимов, которые учитывали бы как высокую степень автономии систем, так и их инструментальную природу, исключающую наличие собственных интересов.

Особую значимость приобретает разработка механизмов распределения бремени доказывания в делах, связанных с ущербом от действий ИИ. В условиях эпистемологической непрозрачности алгоритмов возложение обязанности доказывать причинно-следственную связь на потерпевшую сторону фактически лишает ее возможности судебной защиты. Эволюция доктрины должна двигаться в направлении презумпции причинности или внедрения концепции строгой ответственности владельцев и разработчиков систем за действия своих созданий, аналогично ответственности за источник повышенной опасности, но с учетом информационной специфики. При этом важно избежать ситуации, когда чрезмерное регуляторное давление приведет к стагнации технологического развития; баланс может быть найден через развитие институтов алгоритмического аудита и сертификации, которые станут превентивными фильтрами, не допускающими на рынок потенциально опасные архитектуры.

Социогуманитарный взгляд на проблему выявляет, что вопрос об ответственности ИИ — это, в сущности, вопрос о том, какое общество мы строим и какую роль в нем отводим человеку. Отказ от исключительной монополии человека на принятие решений требует компенсаторных механизмов, обеспечивающих сохранение человеческого достоинства и права на объяснение решений, затрагивающих жизненно важные интересы личности. Перспектива развития правового регулирования видится не в создании новых запретов, а в формировании среды «доверенного искусственного интеллекта», где технические стандарты прозрачности и интерпретируемости станут неотъемлемой частью правовых требований. Это потребует тесной конвергенции права, этики и компьютерных наук, создания междисциплинарных институтов,

способных оперативно реагировать на появление новых технологических вызовов и адаптировать нормативную ткань общества к реальности сосуществования с автономными интеллектуальными агентами.

Библиография

1. Барышников П.Н. Человек и системы искусственного интеллекта / под ред. ак. РАН В.А. Лекторского. СПб.: Юридический центр, 2022. 328 с.
2. Бахтеев Д.В. Правовой статус систем искусственного интеллекта и модели дифференциации юридической ответственности за причиненный ими вред // Lex Russica (Русский закон). 2024. Т. 77. № 4 (209). С. 9-23.
3. Боуэн Г. Агентный искусственный интеллект: правовые и этические вызовы автономных систем // Journal of Digital Technologies and Law. 2025. Т. 3. № 3. С. 431-445.
4. Глуховский А.С., Дурнев А.Д., Чирва Д.В. Распределенная моральная ответственность в сфере искусственного интеллекта // Этическая мысль. 2024. Т. 24. № 1. С. 129-143.
5. Гуркина К.А. О моделях правового регулирования искусственного интеллекта: задачи и перспективы // Молодой юрист. 2021. № 3. С. 40-47.
6. Жигарева Ю.С., Леднева Т.В. Модели ответственности искусственного интеллекта // Юридическая орбита. 2022. № 2. С. 139-141.
7. Козлов И.Н. Правовые аспекты регулирования искусственного интеллекта как фактор устойчивого развития общества и технологического прогресса // Академический исследовательский журнал. 2025. Т. 3. № 1. С. 209-222.
8. Лолаева А.С., Сакаева К.У. Искусственный интеллект: правовые и этические аспекты // Юридические исследования. 2021. № 8. С. 63-71.
9. Михеев М.Ю., Прокофьев О.В., Савочкин А.Е., Семочкина И.Ю. Обеспечение надежности в жизненном цикле систем искусственного интеллекта ответственного назначения // Надежность и качество сложных систем. 2023. № 3 (43). С. 12-20.
10. Никитенко С.В. Концепции деликтной ответственности за вред, причиненный системами искусственного интеллекта // Вестник экономического правосудия Российской Федерации. 2023. № 1. С. 156-174.
11. Романова А.С. Начала законодательства для автономных систем искусственного интеллекта // Надежность. 2024. Т. 24. № 3. С. 10-17.
12. Сангинзода Д.Ш. Ответственность за действия искусственного интеллекта: правовые, этические и технологические аспекты // Азия и Европа. 2025. № 4 (24). С. 136-145.
13. Сливинский А.Б. Доверие к машинному интеллекту: взаимодействие искусственного интеллекта и человека с автономными системами в военных операциях // Авиационные системы. 2025. № 8. С. 41-53.
14. Тиханычев О.В. О разделении ответственности за ошибки эксплуатации робототехнических систем // Программные системы и вычислительные методы. 2025. № 3. С. 59-71.
15. Хмелевской И.Р., Калашников Н.А. Вопросы ответственности за действия искусственного интеллекта в сфере здравоохранения // Вестник Челябинского государственного университета. Серия: Право. 2024. Т. 9. № 4. С. 10-13.

The Problem of Responsibility and Agency in Autonomous Artificial Intelligence Systems with Distributed Decision-Making: A Social-Philosophical Analysis

Yuliya S. Buzykova

PhD in Pedagogical Sciences,
Associate Professor,
Financial University under the Government of the Russian Federation,
125993, 49, Leningradsky ave., Moscow, Russian Federation;
e-mail: Juliaserg_buz@mail.ru

Buzykova Yu.S.

Abstract

The article is devoted to the philosophical-legal problematization of responsibility and agency in autonomous artificial intelligence systems operating in a distributed decision-making mode in critically significant spheres — from transportation and finance to medicine and military applications. It is shown that the stochastic nature of trainable neural networks, the emergent effects of multi-agent architectures, and the epistemological opacity of the "black box" rupture the classical link between the subject's intention, causality, and the imputation of consequences, giving rise to the phenomenon of responsibility diffusion and the "problem of many hands." The semantic shift in the concept of agency from conscious intentionality to the functional ability to initiate changes in the environment is discussed, as well as the limitations of attempts to formalize moral norms in code and to ascribe to the system the status of an "electronic person" given its lack of interests and susceptibility to sanctions. The potential of hybrid models, which distribute responsibility according to the degree of control at the stages of the technology's lifecycle, is argued, and a regulatory drift towards risk-based liability regimes, insurance, algorithmic auditing, and certification is noted. Additionally, the influence of AI on human autonomy through recommender practices and nudging is analyzed, requiring a rethinking of consent doctrines and the right to an explanation of decisions.

For citation

Buzykova Yu.S. (2025) Problema otvetstvennosti i agentnosti v avtonomnykh sistemakh iskusstvennogo intellekta s raspredelennym prinyatiyem resheniy: sotsial'no-filosofskiy analiz [The Problem of Responsibility and Agency in Autonomous Artificial Intelligence Systems with Distributed Decision-Making: A Social-Philosophical Analysis]. *Kontekst i refleksiya: filosofiya o mire i cheloveke* [Context and Reflection: Philosophy of the World and Human Being], 14 (12A), pp. 116-124. DOI: 10.34670/AR.2025.46.53.021

Keywords

Autonomous artificial intelligence systems, distributed decision-making, risk liability, agency, algorithmic audit, AI ethics, legal regulation.

References

1. Barishnikov, P. N. (2022). Chelovek i sistemy iskusstvennogo intellekta [Man and artificial intelligence systems]. Iuridicheskii Tsentr.
2. Bakhteev, D. V. (2024). Pravovoi status sistem iskusstvennogo intellekta i modeli differentsiatsii iuridicheskoi otvetstvennosti za prichinennyi imi vred [Legal status of artificial intelligence systems and models of differentiation of legal liability for harm caused by them]. Lex Russica (Russkii zakon), 77(4(209)), 9–23.
3. Bowen, G. (2025) Agentnyi iskusstvennyi intellekt: pravovye i eticheskie vyzovy avtonomnykh sistem [Agent artificial intelligence: legal and ethical challenges of autonomous systems]. Journal of Digital Technologies and Law, 3(3), 431–445.
4. Glukhovskii, A. S., Durnev, A. D., & Chirva, D. V. (2024). Raspredelennia moralnaia otvetstvennost v sfere iskusstvennogo intellekta [Distributed moral responsibility in the field of artificial intelligence]. Etisheskaia mysl, 24(1), 129–143.
5. Gurkina, K. A. (2021). O modeliakh pravovogo regulirovaniia iskusstvennogo intellekta: zadachi i perspektivy [On models of legal regulation of artificial intelligence: tasks and prospects]. Molodoi iurist, (3), 40–47.
6. Kozlov, I. N. (2025) Pravovye aspeky regulirovaniia iskusstvennogo intellekta kak faktor ustochivogo razvitiia obshchestva i tekhnologicheskogo progressa [Legal aspects of artificial intelligence regulation as a factor of sustainable development of society and technological progress]. Akademicheskii issledovatel'skii zhurnal, 3(1), 209–222.
7. Lolaeva, A. S., & Sakaeva, K. U. (2021). Iskusstvennyi intellekt: pravovye i eticheskie aspeky [Artificial intelligence: legal and ethical aspects]. Iuridicheskie issledovaniia, (8), 63–71.

8. Mikheev, M. Iu., Prokofev, O. V., Savochnin, A. E., & Semochkina, I. Iu. (2023). Obespechenie nadezhnosti v zhiznennom tsikle sistem iskusstvennogo intellekta otvetstvennogo naznacheniia [Ensuring reliability in the life cycle of artificial intelligence systems for responsible purposes]. Nadezhnost i kachestvo slozhnykh sistem, (3(43)), 12–20.
9. Nikitenko, S. V. (2023). Kontseptsii deliktnoi otvetstvennosti za vred, prichinennyi sistemami iskusstvennogo intellekta [Concepts of delictual liability for harm caused by artificial intelligence systems]. Vestnik ekonomicheskogo pravosudiia Rossiiskoi Federatsii, (1), 156–174.
10. Romanova, A. S. (2024). Nachala zakonodatelstva dlja avtonomnykh sistem iskusstvennogo intellekta [Principles of legislation for autonomous artificial intelligence systems]. Nadezhnost, 24(3), 10–17.
11. Sangingzoda, D. Sh. (2025) Otvetstvennost za deistviia iskusstvennogo intellekta: pravovye, eticheskie i tekhnologicheskie aspekty [Liability for actions of artificial intelligence: legal, ethical and technological aspects]. Azia i Evropa, (4(24)), 136–145.
12. Slivitskii, A. B. (2025) Doverie k mashinnomu intellektu: vzaimodeistvie iskusstvennogo intellekta i cheloveka s avtonomnymi sistemami v voennykh operatsiakh [Trust in machine intelligence: interaction of artificial intelligence and human with autonomous systems in military operations]. Aviationnye sistemy, (8), 41–53.
13. Tikhachev, O. V. (2025) O razdelenii otvetstvennosti za oshibki ekspluatatsii robototekhnicheskikh sistem [On the division of liability for errors in the operation of robotic systems]. Programmnaya sistemy i vychislitelnye metody, (3), 59–71.
14. Khmelevskaia, I. R., & Kalashnikov, N. A. (2024). Voprosy otvetstvennosti za deistviia iskusstvennogo intellekta v sfere zdravookhraneniia [Issues of liability for actions of artificial intelligence in the healthcare sector]. Vestnik Cheliabinskogo gosudarstvennogo universiteta. Seriia: Pravo, 9(4), 10–13.
15. Zhigareva, Iu. S., & Ledneva, T. V. (2022). Modeli otvetstvennosti iskusstvennogo intellekta [Models of liability for artificial intelligence]. Iuridicheskaiia orbita, (2), 139–141.