
ЭТИКА

Ю. Ю. ПЕТРУНИН

ГЕНЕРАТИВНЫЙ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ: PRO & CONTRA? ОБЗОР НАУЧНОЙ КОНФЕРЕНЦИИ В СЕКЦИИ «ЭТИКА И ЦЕННОСТИ ГИИ»*

Статья представляет собой обзор ключевых дискуссий, состоявшихся на секции «Этика и ценности генеративного искусственного интеллекта» в рамках международной научной конференции «Генеративный искусственный интеллект в отраслях экономики и социальной сфере: pro&contra» (октябрь 2025 г.). Актуальность темы обусловлена стремительным распространением технологий ГИИ и их трансформирующим воздействием на все сферы общества, что сопровождается ростом публикационной активности в области этики ИИ. В статье применены методы сравнительно-исторического, наукометрического и кластерного анализа.

*В статье систематизированы материалы выступлений по шести ключевым тематическим блокам/кластерам: 1) онтологический статус ГИИ (инструмент vs агент) и пределы его моральной субъектности; 2) критический анализ современных практик этического регулирования, включая феномены *ethics washing* и *ethics bashing*; 3) трансформация межличностных отношений, доверия и когнитивной экологии человека в условиях взаимодействия с ИИ; 4) влияние ГИИ на научную деятельность и творчество, проблемы авторства и цифрового неравенства; 5) геополитическое измерение этики ИИ, противостояние суверенного и глобалистского подходов; 6) когнитивные риски, включая атрофию самостоятельного мышления и угрозу персонализированных манипуляций (когнитивных атак).*

* **Для цитирования:** Петрунин Ю. Ю. Генеративный искусственный интеллект: pro & contra? Обзор научной конференции в секции «Этика и ценности ГИИ» // *Философия и общество*. 2026. № 1. С. 74–87. DOI: 10.30884/jfio/2026.01.04.

For citation: Petrunin Yu. Yu. The Generative Artificial Intelligence: Pros & Cons?: Review of the Scientific Conference in the Section “Ethics and Values of Generative Artificial Intelligence” // *Filosofiya i obshchestvo = Philosophy and Society*. 2026. No. 1. Pp. 74–87. DOI: 10.30884/jfio/2026.01.04 (in Russian).

Философия и общество, № 1 2026 74–87

DOI: 10.30884/jfio/2026.01.04

Основной вывод работы заключается в том, что этика ГИИ представляет собой динамическое поле напряжений между различными полюсами: глобализмом и суверенитетом, инструментальностью и субъектностью, автоматизацией и творчеством. В качестве наиболее перспективных выделяются подходы, которые отказываются от антропоморфизации ГИИ, интегрируют этику в конкретные профессиональные практики, фокусируются на когнитивной устойчивости человека и признают культурную множественность этических оснований. Конференция продемонстрировала переход дискуссии от абстрактных вопросов к поиску конкретных технических и регуляторных решений, где ключевой задачей становится сохранение человеческой моральной агентности при интеграции технологий, служащих развитию, а не деградации человека.

Ключевые слова: генеративный искусственный интеллект (ГИИ), этика ИИ, философия искусственного интеллекта, регулирование ИИ, когнитивная безопасность, субъектность ИИ, цифровое доверие, *ethics washing*, технологический суверенитет.

This article provides an overview of the key discussions held at the “Ethics and Values of Generative AI” section of the international scientific conference “Generative Artificial Intelligence in Economic and Social Sectors: Pro&Contra” (October 2025). The relevance of this topic is driven by the rapid spread of generative AI technologies and their transformative impact on all spheres of society, which is accompanied by an increase in publication activity in the field of AI ethics.

The authors focus on a multifaceted analysis of the ethical challenges posed by generative artificial intelligence. The article applies methods of comparative-historical, scientometric and cluster analysis. The article systematizes presentations across six key thematic blocks: 1) the ontological status of generative artificial intelligence (tool vs. agent) and the limits of its moral agency; 2) a critical analysis of contemporary ethical regulation practices, including the phenomena of “ethics washing” and “ethics bashing”; 3) transformation of interpersonal relationships, trust, and human cognitive ecology in the context of interaction with AI; 4) impact of the generative AI on scientific activity and creativity, issues of authorship and digital inequality; 5) the geopolitical dimension of AI ethics, the confrontation between sovereign and globalist approaches; 6) cognitive risks, including the atrophy of independent thought and the threat of personalized manipulation (cognitive attacks).

The main conclusion of the work is that the ethics of generative AI represents a dynamic field of tension between various poles: globalism and sovereignty, instrumentalism and subjectivity, automation and creativity. The most promising approaches are those that reject anthropomorphization of generative AI, integrate ethics into specific professional practices, focus on human cognitive resilience, and recognize cultural plurality of ethical foundations. The conference demonstrated a shift in discussion from abstract questions to the search

for concrete technical and regulatory solutions, with the key task being preserving human moral agency while integrating technologies that serve human development, rather than degradation.

Keywords: generative artificial intelligence (GAI), AI ethics, philosophy of artificial intelligence, AI regulation, cognitive security, AI subjectivity, digital trust, ethics washing, technological sovereignty.

Введение

С самого начала исследований в области искусственного интеллекта (ИИ) ученые высказывали опасения по поводу этических последствий его использования в обществе [Morley *et al.* 2023]. Более того, этика искусственного интеллекта (ИИ) возникла на 14 лет раньше, чем появился сам термин ИИ. Еще до Дартмутской конференции 1956 г., до статьи Алана Тьюринга [Turing 1950] и книги Норберта Винера [Wiener 1954; Винер 2001] в рассказах Айзека Азимова были сформулированы три этических правила поведения для роботов. Тем не менее до недавнего времени данная тема не входила в топовые. Изменения начались с появлением генеративного искусственного интеллекта (ГИИ). Это хорошо видно по публикационной активности научного сообщества за последние годы (Рис.).

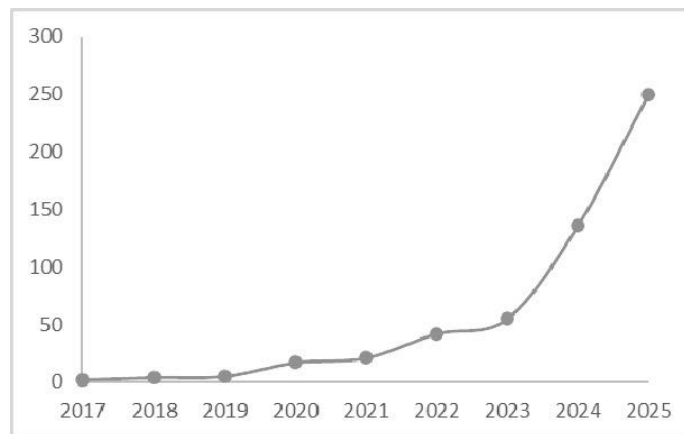


Рис. Динамика публикаций в РИНЦ по теме «Этика ИИ» (на 3 ноября 2025 г.). График построен автором

Стоит отметить, что этику ИИ изучает не только философия. Существенный вклад в данную тематику сделали поведенческая экономика, социология профессий, кибернетика и другие науки. За по-

следние несколько лет опубликовано много важных работ, формирующих междисциплинарную школу, объединяющую технические и гуманитарные подходы [Карпов 2020; Lisanyuk, Petrova 2025; Алексеев, Гарбук 2022; Гарбук 2024; Петруня 2020; Карпов, Королева 2022; Сорокоумов, Карпов 2021; Шахнов и др. 2024], социально-экономические подходы [Gomes *et al.* 2025; Ивахненко, Ковальзон 2024; 2025; Wang *et al.* 2024; Afroogh *et al.* 2024; Huynh, Aichner 2025; Кондрашов и др. 2025; Ryan 2020; Dahlin 2025; Тиссен 2020; Morley *et al.* 2023; Knight *et al.* 2025; Петрунин и др. 2025; Res-séguier, Rodrigues 2020; Paryshev, Chan 2025; Алексеев, Алексеева 2024] и философские подходы [Углева и др. 2024; Флейшер, Шульц 2019; Knight *et al.* 2025; Schultz *et al.* 2025; Разин 2024а; 2024б; Алексеева 2024; Леушина, Карпов 2022].

В Табл. 1 ранжированы по числу статей академические дисциплины, в той или иной степени занимающиеся этикой ИИ. Поскольку количество журналов в РИНЦ по наукам сильно различается, была произведена нормализация статей (количество статей по этике ИИ разделено на количество журналов в данной науке). В итоге по публикационной активности философия заняла первое место в обсуждаемой тематике, на втором месте кибернетика (Табл.).

Табл. Распределение наук в РИНЦ по вкладу в тему «Этика ИИ» (на 3 ноября 2025 г.). Вычислено и ранжировано автором

Наука/тема	Кол-во статей	Кол-во журналов	Нормализация
Экономика	95	449	0,21
Философия	82	113	0,73
Педагогика	47	333	0,14
Юридические науки	41	386	0,11
Кибернетика	24	63	0,38
Автоматика. Вычислительная техника	10	134	0,07
Медицина и здравоохранение	9	594	0,02
Социология	9	148	0,06
Политические науки	7	129	0,05

Рост научных публикаций неудивителен – скорость влияния ГИИ на общество также стремительно растет. Если ежемесячные посещения веб-сайта ChatGPT в январе 2023 г. равнялись 616 млн,

то в сентябре 2025 они достигли 5 млрд 800 млн [Singh 2025]. Однако воздействие ГИИ на общество неоднозначно. К позитивным последствиям можно отнести, например, Нобелевские премии 2024 г. по физике и по химии, которые без искусственных нейросетей и машинного обучения были бы невозможны. Но отметим, что один из лауреатов Нобелевской премии, достигший выдающихся научных результатов с помощью искусственного помощника, Джеффри Хинтон, подписал письмо ведущих ученых и общественных деятелей, призывающее запретить гонку за создание суперинтеллекта [Statement...]. Очевидно, что ГИИ оказывает и негативное воздействие. Проведенные американским исследовательским центром Pew Research Center летом 2025 г. опросы населения показали: 53 % взрослых американцев уверены, что ГИИ ухудшит способность людей мыслить творчески, а 50 % – что он ухудшит способность людей формировать значимые отношения [Kennedy *et al.* 2025]. Более половины американцев (57 %) оценили социальные риски ИИ как высокие по сравнению с 25 %, утверждающими, что преимущества ИИ высоки (Там же).

1 и 2 октября 2025 г. в Институте философии РАН прошло заседание секции «Этика и ценности генеративного искусственного интеллекта» международной конференции «Генеративный искусственный интеллект в отраслях экономики и социальной сфере: pro&contra». Организаторами секции были факультет государственного управления МГУ имени М. В. Ломоносова, Институт философии РАН и Научный совет при Президиуме РАН по методологии искусственного интеллекта и когнитивных исследований. Большинство выступающих являлись авторами работ, указанных выше. Научное мероприятие охватило широкий спектр тем – от фундаментальных философских вопросов о субъектности и моральной автономии искусственных нейросетей до прикладных проблем этического регулирования, ценности места ИИ в образовании, научной практике и новых вызовов когнитивной безопасности.

С приветственным словом выступил академик РАН, председатель Научного совета при Президиуме РАН по методологии искусственного интеллекта и когнитивных исследований В. А. Лекторский. Он напомнил, что 20 лет назад, когда создавался этот совет, мало кто интересовался искусственным интеллектом. Сейчас же о нем говорят все на свете. Безусловно, отметил В. А. Лекторский, будущее человека и человечества связано с развитием ИИ. Но с этим развитием возникают также и проблемы. Как понимать ГИИ? Неко-

торые называют его «сократическим попугаем», неспособным к творчеству и научному познанию. Так ли это? Что может машина, а чего не может? Как нам жить дальше? Философия становится главной наукой для разрешения этих вопросов. «С большим интересом жду сегодняшних дискуссий. Желаю успехов и результатов», – закончил свое выступление В. А. Лекторский.

От Института философии РАН выступила заместитель директора по научной работе О. В. Попова. Она также отметила, что теперь ИИ – часть реальности. ГИИ обладает колоссальной силой, преобразующей общество. Он же ставит фундаментальные вопросы, в исследовании которых философия играет очень важную роль. «Как направить титаническую мощь техники на благо человека?» – этим вопросом закончила свое выступление О. В. Попова.

Дискуссия

Можно выделить несколько наиболее актуальных тем, обсуждавшихся на секции «Этика и ценности ГИИ».

1. Онтология и субъектность ИИ: инструмент или агент?

Центральный философский вопрос, обсуждавшийся на секции, – определение статуса ГИИ: является ли он субъектом или инструментом? И. Ю. Алексеева в своем выступлении «Применение норм служебной этики к генеративному ИИ» предложила важное концептуальное различие: ГИИ – это агент-инструмент, в отличие от человека как агента-субъекта. Это позволяет избежать антропоморфизации и сохранить ответственность за действия ГИИ за человеком, но при этом признать, что к ГИИ применимы некоторые нормы служебной этики. Подобный подход сочетает прагматизм и этическую строгость.

В то же время доклад В. Э. Войцеховича «Категорический императив И. Канта и генеративный ИИ» и доклад Е. Н. Лисанюк и Н. В. Перовой «Логические аспекты аналогии между цифровым правосудием и цифровой моралью» подчеркнули принципиальные ограничения ГИИ: он не способен к творческому моральному суждению, не понимает актуальной бесконечности и не может строить универсальные этические законы. Мораль ГИИ – всегда редуцированная, формализованная и неполная. Это подтверждается логическими парадоксами: этические системы, в отличие от правовых, принципиально неполны и не поддаются алгоритмизации.

2. Этическое регулирование: между декларациями и практикой

Доклад Ю. Ю. Петрунина «Этическое регулирование ИИ и ГИИ» и доклад Д. А. Томильцева «“Ловушки” для гуманитария: феномены “ethics washing” и “ethics bashing” в контексте этики в сфере генеративного искусственного интеллекта» критически проанализировали современные практики этического регулирования. Были выявлены два ключевых феномена:

– Ethics washing – декларативное соблюдение этических норм без реальных механизмов контроля, используемое для избежания жесткого правового регулирования;

– Ethics bashing – обесценивание этики как инструмента, когда ее нормы не обеспечивают прозрачных и действенных решений.

Эти феномены свидетельствуют о кризисе доверия к этическим кодексам. В докладе Н. Н. Опариной «Кодекс ИИ-этики консультантов по управлению» был предложен выход – разработка профессиональных этических кодексов (на примере консультантов по управлению), где этика интегрирована в конкретные практики и ответственность.

Однако В. Э. Карпов в выступлении «О некоторых аспектах применения систем генеративного искусственного интеллекта в задаче формирования цифрового доверия» предупредил: этические декларации могут использоваться для формирования цифрового доверия как инструмента управления обществом, включая мифологизацию ИИ, «формирование духовного и мировоззренческого обоснования веры в Цифру» и замены критического мышления псевдокритическим.

3. Человек и ИИ: диалог, доверие, когнитивная экология

В докладе В. Г. Богомякова и Е. Г. Беляковой «Диалогическая педагогика и этика взаимодействия: анализ концепции “Я и Ты” Мартина Бубера в контексте отношений человека и генеративного искусственного интеллекта в образовании» поднят вопрос о природе взаимодействия с ГИИ в образовании. Авторы предостерегают от редукции педагогического диалога до инструментального общения «Я – Оно», но признают, что граница между «Оно» и «Ты» размывается. Ключевой задачей становится формирование цифровой грамотности: обучение распознаванию границ возможностей ГИИ и критической оценке его ответов.

Доклад Н. З. Нуралиевой «Практики использования технологий генеративного искусственного интеллекта как зеркало доверия к институтам» и доклад Р. Р. Карнеева «От статики к циркуляции: человек как среда индивидуации цифровых субъектов» расширили эту тему: практики использования ГИИ выступают зеркалом доверия к институтам. Если доверие к государству или медицине низкое, пользователь может предпочесть ИИ как «нейтрального» посредника. В случае с цифровыми двойниками умерших (танаботами) человек становится средой индивидуации цифрового субъекта – не просто пользователем, а участником процесса становления новой онтологической сущности.

Т. И. Синюгина и Л. Т. Шароян, основываясь на собственном опыте работы с детьми, страдающими аутизмом, пошли еще дальше. Эти участницы конференции предложили концепцию взаимного развития: при определенном уровне взаимодействия человек и ИИ вступают в режим «восьмерки, превращающейся в ноль» – сотворчества, где оба участника трансформируются и нейросеть становится не инструментом, а партнером в познавательном процессе.

4. Наука, творчество и авторство в эпоху ГИИ

Доклады В. Ю. Перова «Системы искусственного интеллекта как зеркало этических проблем в современной науке», А. В. Разина «Генеративный искусственный интеллект как инструмент научного исследования» и О. Э. Петруни «Генеративный искусственный интеллект: аксиологические аспекты “игры в имитацию”» были посвящены трансформации научной и творческой деятельности в эпоху ГИИ. Новые модели ИИ ускоряют решение рутинных процессов, но и создают следующие риски:

- галлюцинации и каскад ошибок, особенно в графике;
- эрозия авторства: кто автор – пользователь, разработчик или модель?
- подмена научного скептицизма автоматизированной генерацией;
- усиление цифрового неравенства (эффект Матфея).

А. В. Разин в докладе «Генеративный искусственный интеллект как инструмент научного исследования» предложил конструктивный путь: не запрещать использование ГИИ, а обязательно декларировать его применение в методической части работ. Это позволяет сохранить прозрачность и развивать новую этику научной практики.

О. Э. Петруня в докладе «Генеративный искусственный интеллект: аксиологические аспекты “игры в имитацию”» напоминает, что ГИИ реализует тьюринговскую «игру в имитацию», чья суть – введение в заблуждение, поскольку правил игры просто нет. Это ставит под сомнение саму возможность «честного» взаимодействия и «подогревает» мифологию трансгуманизма.

5. Геополитика и цивилизационные ценности

Доклад А. В. Черняева «Этика искусственного интеллекта: суверенный vs глобалистский подход» представил суверенный подход к этике ГИИ, противопоставляя его глобалистской повестке ЮНЕСКО. Автор утверждал, что международные этические рекомендации несут в себе идеологическую нагрузку (неолиберализм, гендерная идеология, мультикультурализм), угрожающую традиционным ценностям России. Автор предложил опираться на русское философское наследие для выработки собственной этической парадигмы.

Доклад Г. О. Куранова и Л. А. Стрижковой «Искусственный интеллект: риски и возможности» развивал эту мысль в практическом ключе: создание национальных больших лингвистических моделей (LLM) на русскоязычной базе может стать инструментом не только технологического суверенитета, но и культурной консолидации («русский мир»).

6. Когнитивные риски и защита сознания

Е. Н. Ивахненко в докладе «Нейросетевая информационная избыточность: pro et contra» и А. В. Углева в докладе «Взлом мозга: защита от когнитивных атак в эпоху искусственного интеллекта» предупредили о когнитивной атрофии: зависимость от ИИ снижает способность к самостоятельному мышлению, памяти, анализу. Это не метафора – нейровизуализация показывает снижение активности мозга у постоянных пользователей.

А. В. Углева ввела понятие когнитивной атаки как персонализированного манипулирования через ГИИ на основе больших данных (BD – Big Data). Защита от таких атак требует не только цифровой гигиены, но и соответствующего образования, формирующего эмоциональное присвоение этических норм.

Заключение: этика как поле битвы и диалога

Конференция показала, что этика генеративного искусственного интеллекта – это не набор правил, а динамическое поле напряжений между такими полюсами, как:

- глобализм и суверенитет,
- инструментальность и субъектность,
- автоматизация и творчество,
- доверие и контроль,
- декларации и практика.

В качестве наиболее перспективных подходов можно выделить те, которые:

- 1) отказываются от антропоморфизации, но признают ГИИ как агент-инструмент;
- 2) интегрируют этику в профессиональные практики, а не ограничиваются абстрактными принципами;
- 3) фокусируются на образовании и когнитивной устойчивости человека;
- 4) признают культурную и цивилизационную множественность этических оснований.

Конференция продемонстрировала, что этика ИИ переходит от абстрактных дискуссий к конкретным техническим и регуляторным решениям. Ключевой вызов – найти баланс между инновациями и защитой человеческой автономии, между глобальными стандартами и национальным суверенитетом в цифровой сфере.

Наблюдается сдвиг от вопросов «можно ли» к вопросам, «как именно» внедрять этические принципы в архитектуру ИИ-систем, что делает дискуссию не только философски глубокой, но и практически значимой. Этика ГИИ – это не просто регулирование технологий, а проекция наших собственных ценностей, страхов и надежд. Как справедливо заметила А. В. Углева, «идея закодировать человеческие ценности в машины... может основываться лишь на неверном предположении о существовании морального прогресса». Поэтому главная задача состоит в том, чтобы сохранить нашу собственную моральную агентность.

Все более широкое применение систем генеративного искусственного интеллекта в профессиональной деятельности, образовании и повседневной жизни требует особого внимания к вопросам этики и ценностей. Общее предназначение искусственной нейросе-

ти, сколь бы высокой степенью автономности она ни обладала, заключается в том, чтобы служить человеку, содействовать развитию человека, а не деградации его интеллектуальных способностей и нравственности. Механизмы этического регулирования должны разрабатываться с учетом специфики областей применения систем генеративного ИИ, правовой системы государства и основополагающих ценностей общества.

Конференция отразила многогранную и междисциплинарную природу современных дискуссий об этике и ценностях искусственного интеллекта, в особенности генеративного ИИ. В совокупности 20 докладов сформировали сложную, дискуссионную, но целостную картину вызовов, парадоксов и перспектив, стоящих перед обществом, наукой, технологиями и философией в эпоху цифровой трансформации и генеративного искусственного интеллекта.

Литература

Алексеев А. П., Алексеева И. Ю. Статус этических кодексов в этике искусственного интеллекта // Информационное общество. 2024. № 4. С. 43–49.

Алексеев А. Ю., Гарбук С. В. Как можно доверять системам искусственного интеллекта? Объективные, субъективные и интересубъективные параметры доверия [Электронный ресурс] : Искусственные общества. 2022. Т. 17. Вып. 2. URL: <https://artsoc.jes.su/s207751800020550-4-1/>. DOI: 10.18254/S207751800020550-4.

Алексеева И. Ю. Этика искусственного интеллекта как прикладная этика // Философия и общество. 2024. № 3(112). С. 69–85.

Винер Н. Человеческое использование человеческих существ: кибернетика и общество / Н. Винер // Человек управляющий. СПб. : Питер, 2001. С. 3–196.

Гарбук С. В. Модель доверия к прикладным системам искусственного интеллекта // Вестник Московского университета. Сер. 21. Управление (государство и общество). 2024. Т. 21. № 4. С. 151–169. DOI: 10.55959/MSU2073-2643-21-2024-4-151-169.

Ивахненко Е. Н., Ковальзон М. М. Будет ли присуждена искусственному интеллекту Нобелевская премия в 2040 году? // Информационное общество. 2024. № 2. С. 2–10.

Ивахненко Е. Н., Ковальзон М. М. Креативность и избыточность информации: угроза выученной беспомощности или творческий ресурс? // Информационное общество. 2025. № 4. С. 2–11.

Карпов В. Э., Королева М. Н. К вопросу о формализации этики поведения коллаборативного робота // Информационные и математические технологии в науке и управлении. 2022. № 4(28). С. 223–233.

Кондрашов П. Е., Петрунин Ю. Ю., Попова С. С. Регулирование разработки и применения искусственного интеллекта: проблемы и направления развития // Вестник Московского университета. Сер. 21. Управление (государство и общество). 2025. Т. 22. № 1. С. 3–19.

Леушина В. В., Карпов В. Э. Этика искусственного интеллекта в стандартах и рекомендациях // Философия и общество. 2022. № 3(104). С. 124–140.

Петрунин Ю. Ю., Кондрашов П. Е., Попова С. С. Этическое регулирование разработки и применения искусственного интеллекта: проблемы и решения // Информационное общество. 2025. № 5. С. 166–197.

Петруня О. Э. Логико-математические основания имитационной стратегии фон Неймана – Тьюринга и проблема определения нейрокомпьютинга // Нейрокомпьютеры: разработка, применение. 2020. Т. 22. № 4. С. 19–27.

Разин А. В. Заметки по поводу законов робототехники А. Азимова // Философия и общество. 2024а. № 4(113). С. 37–48.

Разин А. В. Нейросети и проблема авторства // Научно-исследовательские исследования. 2024б. № 1. С. 46–53.

Сорокоумов П. С., Карпов В. Э. К вопросу о моральных аспектах адаптивного поведения искусственных агентов // Искусственные общества. 2021. Т. 16. № 2.

Тиссен Г. А. Доверие к технологиям искусственного интеллекта: перспективы и риски // Инновационные исследования. 2020. № 1. С. 45–56.

Углева А. В., Шилова В. А., Карпова Е. А. Индекс «этичности» систем искусственного интеллекта в медицине: от теории к практике // Этическая мысль. 2024. Т. 24. № 1. С. 144–159. DOI: 10.21146/2074-4870-2024-24-1-144-159.

Флейшер Р., Шульц Дж. Этическое регулирование искусственного интеллекта // Философия науки и техники. 2019. № 4. С. 35–46.

Шахнов В. А., Пролетарский А. В., Зинченко Л. А., Казаков В. В., Терехов В. В. К вопросу этики искусственного интеллекта // Нейрокомпьютеры: разработка, применение. 2024. Т. 26. № 5. С. 17–29.

Afroogh S., Akbari A., Malone E., Kargar M., Alambeigi H. Trust in AI: Progress, Challenges, and Future Directions [Электронный ресурс] : Humanities and Social Sciences Communications. 2024. Vol. 11. Art. 1568. URL:

<https://www.nature.com/articles/s41599-024-04044-8>. DOI: 10.1057/s41599-024-04044-8.

Dahlin E. Trust in AI // *AI & Society*. 2025. DOI 10.1007/s00146-025-02429-0.

Gomes S., Lopes J. M., Nogueira E. Anthropomorphism in Artificial Intelligence: A Game-Changer for Brand Marketing // *Future Business Journal*. 2025. Vol. 11. No. 1. DOI: 10.1186/s43093-025-00423-y.

Huynh M.-T., Aichner T. In Generative Artificial Intelligence We Trust: Unpacking Determinants and Outcomes for Cognitive Trust // *AI & Society*. 2025. May 6. DOI: 10.1007/s00146-025-02378-8.

Karpov V. E. Can a Robot Be a Moral Agent? // *Artificial Intelligence. RCAI 2020. Lecture Notes in Computer Science*. Vol. 12412. DOI: 10.1007/978-3-030-59535-7_5.

Kennedy B., Yam E., Kikuchi E., Pula I., Fuentes J. How Americans View AI and Its Impact on People and Society [Электронный ресурс] : Pew Research Center. 2025. September 17. URL: <https://www.pewresearch.org/science/2025/09/17/how-americans-view-ai-and-its-impact-on-people-and-society/> (дата обращения: 03.11.2025).

Knight S., Shibani A., Vincent N. Ethical AI Governance: Mapping a Research Ecosystem // *AI and Ethics*. 2025. No. 5. Pp. 841–862. DOI: 10.1007/s43681-023-00416-z.

Lisanyuk E., Perova N. Five Constraints for E-Morality // *Radioelectronics and Information Technology (USBREIT)*. 2025. Pp. 41–44. DOI: 10.1109/USBREIT65494.2025.11054139.

Morley J., Kinsey L., Elhalal A., Garcia F. Operationalising AI Ethics: Barriers, Enablers and Next Steps // *AI & Society*. 2023. Vol. 38. Pp. 411–423. DOI: 10.1007/s00146-021-01308-8.

Papyshev G., Chan K. J. D. Fugazi Regulation for AI: Strategic Tolerance For Ethics Washing // *AI & Society*. 2025. Vol. 40. Pp. 3475–3486. DOI: 10.1007/s00146-024-02084-x.

Rességuier A., Rodrigues R. AI Ethics Should Not Remain Toothless! A Call to Bring Back the Teeth of Ethics // *Big Data & Society*. 2020. Vol. 7(2). DOI: 10.1177/2053951720942541.

Ryan M. In AI We Trust: Ethics, Artificial Intelligence, and Reliability // *Science and Engineering Ethics*. 2020. Vol. 26. Pp. 2749–2767. DOI: 10.1007/s11948-020-00228-y.

Schultz M. D., Conti L. G., Seele P. Digital Ethicswashing: A Systematic Review and a Process Perception Outcome Framework // *AI and Ethics*. 2025. No. 5. Pp. 805–818. DOI: 10.1007/s43681-024-00430-9.

Singh S. Latest Chat GPT Users Stats 2025 (Growth & Usage Report) [Электронный ресурс] : Demand Sage. 2025. October 7. URL: <https://www.demandsage.com/chatgpt-statistics/> (дата обращения: 03.11.2025).

Statement on Superintelligence [Электронный ресурс]. URL: <https://superintelligence-statement.org> (дата обращения: 03.11.2025).

Turing A. M. Computing Machinery and Intelligence // Mind. 1950. Vol. 59(236). Pp. 433–460.

Wang Y.-F., Chen Y.-C., Chien S.-Y., Wang P.-J. Citizens' Trust in AI-Enabled Government Systems // Information Polity. 2024. Vol. 29. No. 3. Pp. 293–312. DOI: 10.3233/IP-230065.

Wiener N. The Human Use of Human Beings: Cybernetics and Society (Revised). London : Doubleday, Anchor, 1954.