

УДК 004.8:17:316.4

DOI: 10.18413/2408-932X-2026-12-1-2-1

Беляев Д. А.<sup>1</sup>,  
Патаев М. В.<sup>2</sup>

**Искусственный интеллект как социально-политический техногибрид эпохи цифрового капитализма: аналитическое размышление над книгой R. Mühlhoff “The Ethics of AI: Power, Critique, Responsibility” (Bristol University Press, Bristol, UK, 2025)**

<sup>1</sup>Липецкий государственный педагогический университет им. П.П. Семенова-Тян-Шанского, ул. Ленина, д. 42, г. Липецк, 398020, Россия; [dm.a.belyaev@gmail.com](mailto:dm.a.belyaev@gmail.com)

<sup>2</sup>Липецкий государственный педагогический университет им. П.П. Семенова-Тян-Шанского, ул. Ленина, д. 42, г. Липецк, 398020, Россия; [mett.pataev@yandex.ru](mailto:mett.pataev@yandex.ru)

**Аннотация.** Актуальность рассмотрения монографии Райнера Мюльхоффа “The Ethics of AI: Power, Critique, Responsibility” (2025) обусловлена тем, что она переносит дискуссию об ИИ из режима техноутопий и «сценариев судного дня» в плоскость критического анализа уже действующих социотехнических практик. Новизна подхода автора связывается с генеалогической критикой: ИИ трактуется как «человеко-сопровождаемая» система, опирающаяся на распределенный человеческий труд и обратную связь пользователей, а власть проявляется в замкнутых контурах управления поведением, в «запечатанных» интерфейсах и в предиктивных режимах знания. В рецензии реконструируются методологические основания книги и обсуждаются ее ключевые концепты (управляющая и кибернетическая власть, предиктивная приватность, коллективная ответственность), показывающие связь алгоритмов с контролем, дискриминацией и структурным неравенством. Результаты рецензии заключаются в выявлении вклада Мюльхоффа в переопределение предмета этики ИИ как политико-философской критики цифрового капитализма. В выводах подчеркивается значимость работы для философии техники и социальной/политической философии, а также для обсуждения общественной подотчетности и регулирования алгоритмических платформ.

**Ключевые слова:** этика ИИ; цифровой капитализм; критика власти; социотехнические системы; предиктивная приватность; кибернетическая власть; коллективная ответственность

**Для цитирования:** Беляев, Д. А. и Патаев, М. В. (2026), «Искусственный интеллект как социально-политический техногибрид эпохи цифрового капитализма: аналитическое размышление над книгой R. Mühlhoff “The Ethics of AI: Power, Critique, Responsibility” (Bristol University Press, Bristol, UK, 2025)», *Научный результат. Социальные и гуманитарные исследования*, 12(1), 224-232. DOI: 10.18413/2408-932X-2026-12-1-2-1

D. A. Belyaev<sup>1</sup>,  
M. V. Pataev<sup>2</sup>

**Artificial intelligence as a socio-political techno-hybrid of the digital capitalism era: analytical reflection on R. Mühlhoff's book "The Ethics of AI: Power, Critique, Responsibility" (Bristol University Press, Bristol, UK, 2025)**

<sup>1</sup>Semenov-Tyan-Shansky Lipetsk State Pedagogical University,  
42 Lenin St., Lipetsk, 398020, Russia;  
*dm.a.belyaev@gmail.com*

<sup>2</sup>Semenov-Tyan-Shansky Lipetsk State Pedagogical University,  
42 Lenin St., Lipetsk, 398020, Russia;  
*mett.pataev@yandex.ru;*

**Abstract.** This review examines Rainer Mühlhoff's monograph *The Ethics of AI: Power, Critique, Responsibility* (2025), which shifts the debate away from techno-utopian or doomsday scenarios toward a critical analysis of everyday socio-technical practices already at work. The novelty of Mühlhoff's approach lies in a genealogical critique of AI as a "human-aided" system: successful AI depends on distributed human labor and feedback loops, and its power materializes through closed control circuits, sealed interfaces, and predictive modes of knowledge. The review reconstructs the book's methodological foundations and discusses its key concepts – governing and cybernetic power, predictive privacy, and collective responsibility – showing how algorithmic infrastructures connect to manipulation, discrimination, and structural inequality. As a result, the review highlights Mühlhoff's contribution to redefining AI ethics as political-philosophical critique of digital capitalism. The conclusion stresses the book's relevance for philosophy of technology, social and political philosophy, and contemporary discussions of accountability and regulation of algorithmic platforms.

**Keywords:** AI ethics; digital capitalism; critique of power; socio-technical systems; predictive privacy; cybernetic power; collective responsibility

**For citation:** Belyaev, D. A. and Pataev, M. V. (2026), "Artificial intelligence as a socio-political techno-hybrid of the digital capitalism era: analytical reflection on R. Mühlhoff's book 'The Ethics of AI: Power, Critique, Responsibility' (Bristol University Press, Bristol, UK, 2025)", *Research Result. Social Studies and Humanities*, 12(1), 224-232, DOI: 10.18413/2408-932X-2026-12-1-2-1

Рецензируемая монография – "The Ethics of AI: Power, Critique, Responsibility" (2025) («Этика ИИ: власть, критика, ответственность») – представляет собой масштабное философское исследование этических аспектов искусственного интеллекта (ИИ) в контексте власти и общества. Ее автор – немецкий философ и математик Райнер Мюльхофф, доктор философии, профессор этики ИИ в Университете Оснабрюка (Германия). Академическая карьера Мюльхоффа развивалась на стыке континентальной философии, социальных исследований технологий и компьютерных наук. В 2016 г. он защитил диссертацию по философии на тему аффекта у Б. Спинозы и М. Фуко, а в 2021 г. возглавил первую в Германии кафедру этики ИИ. Научные интересы Мюльхоффа лежат в области критической философии цифровой эпохи: он исследует взаимосвязи технологий, власти и субъективности, публикуясь по проблемам приватности, дискриминации и больших данных в цифровом обществе. Такой

исследовательский бэкграунд позволяет Мюльхоффу органично сочетать глубокий философский анализ с пониманием технических основ ИИ.

Актуальность тематики монографии не вызывает сомнений. Проблема этики ИИ сегодня – один из центральных вызовов для социальной философии и философии техники. Стремительное проникновение алгоритмов машинного обучения во все сферы жизни сопровождается новыми формами власти и воздействия на людей. Сегодня мы так или иначе вовлечены в работу систем ИИ и больших данных, становясь соучастниками их глобального влияния. Тем самым этика ИИ перестает быть узкой прикладной дисциплиной и превращается в пространство критического осмысления современности. Монография Мюльхоффа посвящена именно этой широкой проблематике: она исследует ИИ как повседневный социотехнический феномен, меняющий социальные нормы и отношения власти. Работа предлагает взглянуть на ИИ сквозь призму философской критики власти. Как писал еще в прошлом веке Лэнгдон Виннер, технические артефакты содержат в себе политические качества (Winner, 1980: 125); Мюльхофф развивает эту мысль применительно к ИИ, показывая, что алгоритмы и данные встроены во властные отношения и способны усиливать существующие структуры контроля и неравенства. Его работа находится на переднем крае междисциплинарного диалога – на пересечении философии техники, этики и исследований данных (Data Studies).

Методология исследования в книге Мюльхоффа опирается на традицию критической философии и генеалогического метода. Автор прямо заявляет: «Эффективная этика должна проистекать из критики, которая носит системный и саморефлексивный характер» (Mühlhoff, 2025: 12). Под «критикой» исследователь понимает глубокий анализ условий и предпосылок, породивших современную ситуацию с ИИ. Теоретические корни автор выводит от Ф. Ницше и К. Маркса через М. Фуко и постструктурализм, а также феминистскую и постколониальную мысль. Немецкий мыслитель отмечает, что идея «аполитичного или этически нейтрального понятия ИИ (да и технологии в целом) – иллюзия» (Mühlhoff, 2025: 5). Это методологическое убеждение ведет Мюльхоффа к особой двухступенчатой стратегии: сначала – критический разбор власти, заключенной в технологиях ИИ, и лишь затем формулирование нормативных выводов.

Переходя к содержанию монографии, следует отметить, что книга тщательно структурирована и охватывает широкий круг вопросов. Уже во введении Мюльхофф обозначает двойную цель своей работы: с одной стороны, дать научно обоснованный разбор вызовов, связанных с ИИ, а с другой – «активизировать читателя как этического и политического агента», побудить к коллективному действию (Mühlhoff, 2025: 7). Внимание акцентируется на непреднамеренной соучастности пользователей из развитых стран (так называемого «глобального Севера») в деятельности крупных технокорпораций. Автор прямо указывает: «Данные, которые мы – привилегированные пользователи – невольно создаем, питают системы, обеспечивающие контроль, манипуляцию, эксплуатацию и дискриминацию людей во всем мире» (Mühlhoff, 2025: 2). Именно этот тезис задает тон всей книге.

Первая часть книги («Власть ИИ»), состоящая из четырех глав, посвящена переосмыслению самого феномена ИИ и цифровых технологий через призму власти. Здесь Мюльхофф задается вопросом, какое именно явление скрывается под наименованием «искусственный интеллект». Он утверждает, что общественное воображение часто одержимо образом некоего «генерализованного ИИ» – автономного, возможно даже сознательного агента. Этот образ подпитывает фантазии о восстании машин и техноутопии, однако подобные футуристические сценарии ИИ, по его мнению, являются лишь идеологическим приемом, формой отвлечения от насущных проблем. Реальный же ИИ, о котором говорит Мюльхофф, – это преимущественно статистические алгоритмы, «машины предсказания», обученные на огромных массивах данных. Автор показывает, что нынешний успех методов машинного

обучения держится вовсе не на каких-то прорывах в подражании человеческому разуму, а на банальной доступности колоссальных объемов человеко-сгенерированных данных и на привлечении человеческого участия для разметки и корректировки алгоритмов (Mühlhoff, 2025: 33-33). Этот тезис опирается на постмарксистскую идею «пленения человеческого разума», в частности, М. Паскуинелли отмечает, что «внутренний код ИИ – это не имитация биологического интеллекта, а интеллект труда и социальных отношений» (Pasquinelli, 2023: 2). Мюльхофф солидаризируется с этой позицией, полагая, что переход от парадигмы имитации (моделирования разума) к парадигме «экспроприации» человеческих способностей – центральный для понимания природы современного ИИ. Современные ИИ-системы не столько заменяют, сколько используют человеческий интеллект, «встраивая» его как компонент в масштабные сети обработки данных. Эту скрытую зависимость автор характеризует как неосознанный цифровой труд пользователей. Соответственно, в интерпретации Мюльхоффа, ИИ предстает как социотехническая система и составляющая режима современного цифрового капитализма.

Далее немецкий исследователь развивает эту мысль в виде целостной концепции. Он вводит термин «человеко-сопровождаемый ИИ» для описания архитектуры современных ИИ-систем. Нынешний ИИ – это всегда социотехнический гибрид, где машина и человек переплетены. Автор приводит различные примеры, призванные доказать, что ИИ опирается на масштабную рекрутировку человеческого интеллекта, причем зачастую «безвозмездно, скрыто или принудительно» (Mühlhoff, 2025: 39). Мюльхофф сравнивает такую модель с классическими кибернетическими системами управления: алгоритмы действуют по принципу «замкнутого контура», постоянно получая обратную связь от человека и подстраивая под нее свои результаты. Он называет эту форму «управляющей властью». Автором подробно анализируются два ее измерения – манипуляция и дискриминация. Под манипуляцией понимаются способы интерфейсного воздействия, побуждающие людей действовать в интересах системы. Под дискриминацией – способность алгоритмов дифференцированно подходить к разным группам, воспроизводя социальные предрассудки или уязвимости. Мюльхофф указывает, что такие эффекты не являются случайными сбоями – напротив, «риск манипуляции и дискриминации пользователей присущ современному ИИ», поскольку он по самой сути стремится управлять поведением и решениями людей ради коммерческих целей (Mühlhoff, 2025: 52-53).

В третьей главе Мюльхофф обращается к вопросу об интерфейсах и пользовательском опыте как особом измерении власти. Здесь автор предлагает оригинальное понятие «запечатанного интерфейса» для описания современных трендов в дизайне устройств и приложений. Речь идет о том, что все больше цифровых технологий скрывают свою техническую сложность за закрытыми пользовательскими оболочками. Интерфейс становится единственным «лицом» устройства, за которым полностью невидимы внутренние процессы обработки данных. Мюльхофф рассматривает исторический сдвиг от ранних персональных компьютеров (где пользователь мог видеть файловую систему, настройки, «внутренности» программы) к современным гаджетам с упрощенными иконками, автоматическими обновлениями и все меньшей прозрачностью. Он отмечает, что примерно с 2010-х гг. индустриальный дизайн культивирует «натурализованное поведение» устройств – то есть стремится сделать технику похожей на понятные быденные объекты, отучая пользователя от инструментального, осмысленного отношения к ней. «Запечатанная» поверхность экрана уже не предполагает, что пользователь будет заглядывать внутрь системы или модифицировать ее работу – напротив, постулируется, что пользователь якобы «не желает и не должен» этого делать (Mühlhoff, 2025: 57). Эту тенденцию Мюльхофф называет «демобилизацией инструментальности»: современного индивида постепенно отучают воспринимать цифровые устройства как инструменты для достижения собственных целей, предлагая взамен закрытые

системы «под ключ», где все решения принимаются за него (Mühlhoff, 2025: 60-62). Власть дизайна, по Мюльхоффу, проявляется именно в этих тонких ограничениях свободы: когда выбор пользователя направляется интерфейсными «подсказками», когда сложные настройки спрятаны, а вместо них предлагаются упрощенные кнопки – пользователь теряет возможность критического и творческого обращения с техникой. В совокупности цифровое «контрпросвещение», по Мюльхоффу, заключается в том, что технологии, вместо просветительского идеала расширения знания и автономии, начинают действовать противоположно – ограничивать знания пользователя о работе системы и тем самым усиливать власть ее создателей (Mühlhoff, 2025: 72-73). По сути, немецкий исследователь показывает, что борьба за власть проходит сегодня не только на уровне законов или экономики, но и на уровне дизайна интерфейсов, формирующих наше опытное поле.

Наконец, в четвертой главе рассматривается тема влияния ИИ-технологий на формирование субъекта. Здесь Мюльхофф переходит к языку социальной философии: опираясь на работы Фуко, он анализирует, каким образом цифровые практики делают людей «субъектами» – как подчиненными, так и носителями идентичности. В духе Фуко автор напоминает, что власть не просто подавляет, но и производит субъективность (Mühlhoff, 2025: 77). Влияние алгоритмических систем на человека рассматривается через эту призму: когда персональные ленты новостей подстраиваются под поведение пользователя, когда рекомендации и рейтинги постоянно отражают ему его же профиль, индивид начинает видеть себя через категории, навязанные цифровой системой. Мюльхофф проводит параллель с понятием «идеологических аппаратов» Л. Альтюссера. Как школы или СМИ в XX в. воспроизводили господствующую идеологию, «замещая» человеку спонтанное мировосприятие предзаданными рамками, так и сегодня интерфейсы соцсетей и сервисов выполняют схожую функцию, встроенную в капиталистические отношения. Мюльхофф указывает, что эти механизмы поддерживают существующие структуры власти, заставляя индивидов подстраиваться под требуемые модели поведения (потребления, самопрезентации и лояльности к системе). Классические этические подходы исходят из признания автономного индивида, свободно принимающего решения, но в условиях, когда поведение миллионов людей тонко направляется алгоритмами и дизайн-паттернами, такая предпосылка проблематична. Поэтому Мюльхофф предлагает осмыслить этику ИИ как проект, где субъект – не данность, а результат властных отношений, и ответственность должна быть переориентирована с индивидуальных моральных качеств на структурные условия, формирующие эти качества. В конце главы автор вводит понятие «коллективной субъектности»: перед лицом глобальных алгоритмических систем пользователи должны осознать себя частью коллективного «мы», несущего ответственность за условия своего существования (Mühlhoff, 2025: 82).

Вторая часть монографии («Власть предсказания»), состоящая из трех глав, сосредоточена на феномене прогнозирующих алгоритмов и связанных с ними эпистемических и этических проблемах. Мюльхофф подробно разбирает, что современный ИИ, по сути, есть технология предсказания на основе данных. Он указывает, что машинное обучение решает задачи, которые классическая эпистемология относилась скорее к индуктивным предположениям, нежели к дедуктивному знанию (Mühlhoff, 2025: 88). Алгоритмы на основе больших данных делают вероятностные выводы о будущем или неизвестном – будь то прогноз покупательского спроса, вероятность заболевания или склонность человека к определенным действиям. Важное следствие – смещение в характере знания: если традиционная наука искала объяснения и причинные модели, то культура больших данных довольствуется коррелятивными предсказаниями, часто непрозрачными по своей сути. Автор приводит сравнение: там, где ученый объясняет явление через общие принципы, алгоритм просто выдает ожидаемый исход на основании статистической схожести случаев (Mühlhoff, 2025: 96-

97). Это ведет к своеобразному «культурному повороту к предсказанию»: растет доверие к цифрам и моделям, которые ничего не объясняют, но дают полезные утилитарные результаты. С этической точки зрения, как отмечает Мюльхофф, такая ситуация проблематична: во-первых, непрозрачность алгоритмов затрудняет общественный контроль; во-вторых, слепое упование на предиктивные модели чревато новой формой детерминизма – когда людей оценивают и обращаются с ними на основе статистического прогнозирования. Автор подчеркивает: власть современных ИИ – это не только власть над настоящим, но и власть определять будущее, навязывая определенные сценарии на основании прогнозов.

Далее немецкий исследователь раскрывает, почему классические представления о конфиденциальности недостаточны в эпоху предиктивных алгоритмов. Ранее приватность понималась как контроль индивида над своими личными данными – что он сам раскрыл, то и может быть использовано. Однако в эпоху больших данных и ИИ эта модель рушится. Мюльхофф отмечает: «Личные данные больше не ограничены тем, что мы добровольно предоставили – они включают и то, что о нас вычислили» (Mühlhoff, 2025: 106). Он предлагает расширить понятие приватности до «предиктивной приватности» – права не только на защиту данных, но и на защиту от чужих алгоритмических выводов о себе (Mühlhoff, 2025: 112). Анализируются три типа угроз приватности со стороны ИИ: персональная таргетированная реклама, которая использует предсказанные интересы; предиктивные системы правопорядка (например, прогнозирующие правонарушения) и медицинские предиктивные модели, которые могут выявить вероятные болезни до того, как о них узнает сам человек. В каждом случае нарушения приватности носят коллективный характер: модель делает вывод о конкретном индивиде, опираясь на данные множества других людей. Поэтому приватность уже не может пониматься лишь индивидуально. Мюльхофф утверждает, что защита приватности должна сместиться с уровня индивидуальных согласий и настроек на уровень коллективных ограничений на алгоритмы (Mühlhoff, 2025: 115-116). Автор подводит читателя к выводу, что необходима новая этическая парадигма приватности, которая бы учитывала предиктивное измерение и коллективное измерение хранения и анализа данных.

Седьмая глава носит более рефлексивный характер и связывает обсуждение предиктивных технологий с философией знания и культуры. Здесь Мюльхофф задается вопросом: как распространение предсказательных алгоритмов изменяет наше понимание знания, ответственности и справедливости? Он противопоставляет инференциальное знание эпохи Просвещения – стремление объяснять явления, апеллируя к разуму, – современной предиктивной рациональности, удовлетворяющейся корреляциями. В этической перспективе это сдвигает акцент с мотивов и причин на результаты и вероятности, что, в частности, порождает возникновение парадокса «опережающей ответственности» и требует переосмысления многих этических концепций (Mühlhoff, 2025: 122-124). Особое внимание автор уделяет рассмотрению идеи справедливости в контексте алгоритмов, отмечая известные примеры «алгоритмической предвзятости» (Mühlhoff, 2025: 129). По мысли Мюльхоффа, этическая культура алгоритмических предсказаний, помимо технических усовершенствований, требует еще политического и социального вмешательства – изменения подходов к сбору данных, прозрачности и подотчетности алгоритмических решений (Mühlhoff, 2025: 131-132). Резюмируется, что этика ИИ должна работать на стыке эпистемологии (понимая природу алгоритмического знания) и социальной философии (учитывая структурное угнетение и историческую несправедливость).

Третья часть монографии («Власть контроля») фокусируется на обсуждении ИИ как инструмента социального управления и связанных с этим проблем непрозрачности, предвзятости и ответственности. Мюльхофф утверждает, что современные системы ИИ в социальной сфере, по сути, стали масштабными «кибернетическими Левиафанами»: они непрерывно мониторят поведение масс людей и корректируют это поведение посредством

рекомендаций, штрафов и рейтингов. Автор предлагает смотреть на алгоритмы «больших платформ» (социальных сетей, интернет-магазинов, поисковых систем и т. д.) как на замкнутые контуры управления вниманием: они измеряют реакцию пользователей, затем регулируют подачу контента, тем самым снова изменяя поведение пользователей (Mühlhoff, 2025: 139-140). Эта модель сверхмасштабного кибернетического контроля отличает цифровую эпоху от прежних форм власти. Теперь власть предельно динамична, постоянно перенастраивается алгоритмами для удержания желаемого состояния. Мюльхофф обсуждает, не является ли такая система качественно новым видом власти, превосходящим по эффективности традиционные институты. Он вводит понятие «кибернетической власти», предполагая, что слепое применение технологии в социальных процессах может привести к состоянию, когда алгоритмическая обратная связь начинает определять нормы (Mühlhoff, 2025: 142-143). Автор подспудно говорит, что ИИ – кульминация давней мечты власти стать тотальной, но невидимой. При этом он подчеркивает, что кибернетическая власть является порождением определенных социальных интересов и экономических сил (Mühlhoff, 2025: 144). Алгоритмы обслуживают цели корпораций, а значит, их контроль небезличен. Поэтому необходимо, по мысли Мюльхоффа, вернуть обсуждение ИИ из сферы узко технической в сферу политической экономии, ставя вопросы, кто контролирует эти контуры обратной связи и в чьих интересах они настроены?

Далее немецкий исследователь обращается к проблеме «черного ящика» ИИ и подходит к вопросу с неожиданной стороны. Мюльхофф в духе своей критической позиции спрашивает: кому выгодна непрозрачность? В целом, непрозрачность становится формой власти, когда люди не могут оспорить решение машины, потому что не понимают, как оно принято. Соответственно, объяснимость ИИ выступает не просто технической задачей, а требованием справедливости. Автор настаивает, что требуются юридические механизмы, обязывающие раскрывать основы алгоритмических решений, особенно в сферах высокого риска. Мюльхофф резюмирует: «Чем больше власть алгоритма над людьми, тем выше должна быть его прозрачность и подконтрольность обществу» (Mühlhoff, 2025: 154).

В последней главе Мюльхофф развивает понятие «коллективной ответственности» применительно к ИИ. Последняя, по его мысли, заключается в «осознании широко разделяемого поведения, делающего возможными те или иные риски» (Mühlhoff, 2025: 163). Осознание общей уязвимости и соучастия может стать основой нового типа солидарности. В его рамках должно прийти понимание, что только коллективным действием можно изменить курс развития ИИ.

Заключение книги оформлено как «Манифест...», что подчеркивает практическую направленность выводов. Этот манифест содержит несколько тезисных пунктов, синтезирующих основные идеи. В частности, Мюльхофф провозглашает необходимость: (1) ставить критику власти на первое место при обсуждении ИИ (т. е. разоблачать реальные социальные условия, стоящие за технологиями, прежде чем говорить об абстрактной морали); (2) рассматривать ИИ как систему, встроенную в человеческий труд и жизнь; (3) коллективизировать ответственность – признать, что этические проблемы ИИ нельзя решить усилиями отдельных «этичных программистов» или осознанных пользователей, нужна политическая воля и совместные действия институтов и граждан; (4) принять новые понятия и ценности, такие как предиктивная приватность, коллективное благо, защита уязвимых групп, которые выходят за рамки классического либерального фокуса на индивидуальной свободе; (5) ужесточить регулирование технокомпаний, связанных с ИИ, законодательно ограничить наиболее опасные практики. По сути, манифест Мюльхоффа – это призыв к «этике сопротивления», которая должна противостоять злоупотреблениям, мобилизуя общество и государство. Автор ясно указывает, что без коллективной

политической борьбы этика останется беспомощной перед лицом глобальных игроков цифрового капитализма.

Подводя итог, можно сказать, что монография Райнера Мюльхоффа произвела достаточно сильное впечатление своей цельностью, глубиной и актуальностью. К сильным сторонам книги прежде всего относится ее междисциплинарность: автор свободно соединяет философский анализ с конкретными технокультурными примерами. Это позволяет читателю увидеть «большую картину» – как исторические идеи и структуры власти продолжают действовать через новейшие цифровые инструменты. Мюльхофф демонстрирует высокую эрудицию: он знаком с последними работами по этике ИИ, упоминает результаты исследований в области data science, права и медиатеории. При этом исследование остается доступным по стилю и языку. Можно отметить удачное введение новых терминов («предиктивная приватность», «запечатанные интерфейсы», «человеко-сопровожаемый ИИ»): они хорошо пояснены и подкреплены примерами, что облегчает их понимание. Еще один несомненный плюс – цельность аргументации, которая подкрепляется повторяющимися рефренами (о соучастности, о структурном вреде, о коллективной ответственности), благодаря чему читатель не теряет нить рассуждений. Монография подкупает и социальной значимостью: Мюльхофф не прячется за академическую отстраненность, он прямо говорит о современных вызовах – будь то дискуссия вокруг данных ChatGPT или опасность праворадикального использования ИИ. Такая откровенная ангажированность философа придает работе особую актуальность.

Разумеется, при всех своих достоинствах, книга не лишена и спорных моментов, открывающих поле для дискуссии. Так, в монографии чувствуется определенная идеалистическая установка на революционный поворот – автор апеллирует к «политическому пробуждению» масс, что само по себе похвально, но оставляет вопрос: насколько достижима эта утопия? Другой возможный слабый момент – фокус на глобальном Севере. Мюльхофф откровенно пишет, что рассматривает ситуацию с точки зрения белого европейского пользователя. С одной стороны, это честно и оправдано, но с другой – некоторые проблемы глобального Юга остаются несколько в тени. Возможно, включение большего числа кейсов из небогатых стран сделало бы анализ еще более убедительным и всеобъемлющим. Наконец, кто-то может упрекнуть автора в отсутствии четких решений: призывы к коллективной ответственности и регулированию звучат вдохновляюще, но практические шаги очерчены лишь в общем. Впрочем, едва ли это можно считать серьезным изъяном, т. к. цель книги все-таки скорее пробудить критическое сознание, нежели дать конкретную программу действий по этизации ИИ.

В целом книга Райнера Мюльхоффа заставляет переосмыслить привычные подходы: вместо того чтобы спрашивать: «Как запрограммировать ИИ быть этичным?», читатель начинает задаваться вопросом: «Каким должен быть наш социотехнический строй, чтобы ИИ служил обществу, а не углублял несправедливость?». Исследование существенно обогащает философский дискурс об этике ИИ, переводя его из модуса абстрактных принципов в плоскость анализа власти, структурной критики и коллективной ответственности.

#### Литература References

- Mühlhoff, R. (2025), *The Ethics of AI: Power, Critique, Responsibility*, Bristol University Press, Bristol, UK. DOI: 10.51952/9781529249262
- Pasquinelli, M. (2023), *The Eye of the Master: A Social History of Artificial Intelligence*, Verso Books, London, UK.
- Winner, L. (1980), “Do artifacts have politics?”, *Daedalus*, 109(1), 121-136.

*Информация о конфликте интересов: авторы не имеют конфликта интересов для деклараций.  
Conflict of Interests: the authors have no conflict of interests to declare.*

**ОБ АВТОРАХ:**

**Беляев Дмитрий Анатольевич**, доктор философских наук, профессор кафедры философии, политологии и теологии, Липецкий государственный педагогический университет им. П.П. Семенова-Тян-Шанского, ул. Ленина, д. 42, г. Липецк, 398020, Россия; [dm.a.belyaev@gmail.com](mailto:dm.a.belyaev@gmail.com)

**Патаев Матвей Валентинович**, аспирант кафедры философии, политологии и теологии, Липецкий государственный педагогический университет им. П.П. Семенова-Тян-Шанского, ул. Ленина, д. 42, г. Липецк, 398020, Россия; [mett.pataev@yandex.ru](mailto:mett.pataev@yandex.ru);

**ABOUT THE AUTHORS:**

**Dmitry A. Belyaev**, Doctor of Philosophy, Professor, Department of Philosophy, Political Science and Theology, Semenov-Tyan-Shansky Lipetsk State Pedagogical University, 42 Lenin St., Lipetsk, 398020, Russia; [dm.a.belyaev@gmail.com](mailto:dm.a.belyaev@gmail.com)

**Matvey V. Pataev**, Postgraduate Student, Department of Philosophy, Political Science and Theology, Semenov-Tyan-Shansky Lipetsk State Pedagogical University, 42 Lenin St., Lipetsk, 398020, Russia; [mett.pataev@yandex.ru](mailto:mett.pataev@yandex.ru);