

ОТЗЫВ

официального оппонента, доктора философских наук, профессора кафедры философии ФГБОУ ВО «Московский авиационный институт (национальный исследовательский университет)» – Цвык Ирины Вячеславовны – о диссертационной работе Бадмаевой Майны Харлановны на тему «Социально-философские проблемы и принципы применения систем искусственного интеллекта», представленной на соискание ученой степени кандидата философских наук по специальности 5.7.7. – Социальная и политическая философия (философские науки)

1. Актуальность темы диссертационного исследования

Актуальность темы диссертационного исследования М.Х. Бадмаевой не вызывает сомнений и обусловлена рядом факторов. Во-первых, очевидно, что сегодня мир переживает период стремительного развития искусственного интеллекта (ИИ). Достижения в этой области приводят к появлению машин, способных учиться и выполнять когнитивные задачи, с которыми раньше могли справиться только люди. Такое технологическое развитие, по-видимому, будет иметь далеко идущие социальные и культурные последствия, что существенно актуализирует социально-философские проблемы развития и применения систем ИИ.

Во-вторых, поскольку искусственный интеллект является когнитивной технологией, различные ее аспекты неразрывно связаны с основными формами деятельности людей: образованием, наукой, культурой и коммуникацией. В области научных исследований ИИ занимает ведущие позиции в анализе и интерпретации данных. В дополнение к этому продолжающееся замещение труда людей интеллектуальными технологиями требует новых инструментов обеспечения жизнеспособности и гибкости трудовых ресурсов. Все чаще современные мыслители высказывают опасения о том, что ИИ может создать угрозу для существования человечества, поскольку он способен взять под контроль многие аспекты нашей повседневной жизни и организации общественной сферы. Таким образом, бесспорна актуальность и насущная необходимость социально-философского анализа и истолкования процесса все возрастающего влияния систем ИИ на жизнь современного общества, негативных последствий его применения, поиск и выработка вектора конструирования будущего человечества.

Наконец, нужно с сожалением констатировать, что мировоззренческое осмысление происходящих изменений все еще существенно отстает от темпов внедрения технологии ИИ, оставляя без ответа многочисленные вызовы, брошенные человеческой цивилизацией «умными» машинами. В этом контексте наблюдается очевидный дефицит социально-философских исследований, спо-

собных аргументированно подойти к решению проблем ИИ с позиции взвешенного рационального анализа, основанного на гуманистических ценностях, лишённого в равной степени откровенно алармистских настроений и ничем не оправданных исключительно радужных, позитивных ожиданий от быстрого прогресса систем ИИ.

Диссертационное исследование М. Х. Бадмаевой, посвященное исследованию социально-философских проблем и принципов применения систем ИИ, на наш взгляд, способно внести значимый научно-теоретический и практический вклад в решение вышеозначенных проблем.

2. *Обоснованность научных положений, выводов и рекомендаций, сформулированных в диссертации.*

Сформулированные в работе научные положения о сущности ИИ, областях его присутствия в жизни общества, специфических проблемах, ставших следствием его применения и философских принципах, обеспечивающих безопасность и эффективность взаимодействия человека с системами ИИ, основаны на значительном по объёму и разнообразию комплексе источников, отражающих современный уровень возможностей ИИ и международный опыт использования систем ИИ в различных сферах общественной жизни. Привлечение в качестве эмпирической базы исследования широкого круга документов, созданных государственными структурами, коммерческими компаниями и общественными организациями, регламентирующих этико-правовые проблемы ИИ, обеспечило необходимый уровень соответствия полученных соискателем результатов требованию обоснованности положений, выводов и рекомендаций, сформулированному Положением ВАК о присуждении ученых степеней. М.Х. Бадмаева впервые использовала в своей диссертации и ввела тем самым в научный оборот около ста документов, позволяющих составить достаточно полное представление об уровне и масштабах применения ИИ в социальной практике и степени развития негативных следствий его активного внедрения в жизнь современного социума.

Приведенные и тщательно проработанные документы мы оцениваем как в полной мере репрезентативный материал, позволяющий решить поставленные автором задачи, связанные с выявлением основных проблем ИИ, а также смысла и значения философских принципов, призванных обеспечить эффективность и безопасность в отношениях человека и ИИ.

3. *Достоверность, обоснованность и новизна полученных выводов*

Достоверность результатов диссертационного исследования определяется серьезной теоретической базой, представленной трудами отечественных и зарубежных ученых, которая позволила сформировать понятийный аппарат диссертации, выявить наиболее острые проблемы, тенденции и закономерности развития ИИ, обусловленные эволюцией и расширением сферы применения ИИ, определить основные пути их преодоления. Достоверность также обеспечивается значительным количеством включенных в исследование документов этического и правового характера, содержащих актуальные данные

о результатах использования ИИ. Общий объем изученной литературы составляет 253 источника, включая исследования зарубежных авторов.

Обоснованность результатов исследования определяется репрезентативностью эмпирических данных, корректным и релевантным применением современных и хорошо зарекомендовавших себя в социально-философском дискурсе приемов и методов научного исследования (системного, деятельностного подхода, герменевтического метода исследования и принципа историзма, традиционных общенаучных методов (анализ, синтез, дедукция, индукция и др.).

Диссертационное исследование содержит целый ряд моментов, которые, несомненно, обладают **научной новизной**. Среди них, в качестве наиболее существенных, можно выделить следующие:

- исследованы и систематизированы документы, регулирующие этические правовые аспекты разработки, внедрения и применения ИИ, содержание которых позволяет выявить основные риски и угрозы для безопасного и эффективного ИИ;

- обоснована необходимость социально-философского исследования проблем применения систем ИИ, позволяющего подойти к их решению с позиций целостности социальной жизни и присущей философии общества нацеленности на индивида, его многогранные потребности и стремление к достижению социального идеала;

- определены социальные проблемы, возникающие вследствие применения систем ИИ, отражающие основные проявления отрицательного влияния ИИ на человека: проблемы причинения вреда, социальной несправедливости и нарушения автономии человека.

- в отдельную группу выделены проблемы, указывающие на причины возникновения нежелательных для человека и общества последствий взаимодействия с ИИ: проблемы непрозрачности, отсутствия ответственности и нарушения конфиденциальности.

- раскрыто содержание философских принципов прозрачности, ответственности, конфиденциальности, представляющих собой условия безопасного и эффективного использования ИИ в различных областях жизни современного общества;

- установлен смысл требований, составляющих содержание принципов социальной справедливости, автономии человека и непричинения вреда, представляющих отдельные стороны, составляющие безопасного, эффективного применения систем ИИ в различных областях жизни социума, а также обосновано основополагающее значение принципа непричинения вреда во всей совокупности рассмотренных выше философских принципов.

К числу бесспорных достоинств диссертации следует отнести реализацию в качестве методологической основы диссертации логического принципа тождества: все понятия, используемые в исследовании четко определены, установлены их предметные поля. В первую очередь это относится к самому

понятию «искусственный интеллект», которое в современной исследовательской литературе, как справедливо отмечает автор, часто носит скорее метафорический, чем предметный характер. Уже в первом параграфе диссертации автор четко определяет, что в рамках данного исследования анализируется так называемый «слабый» искусственный интеллект, это четко очерчивает предметные рамки работы.

Одной из важных новаций работы является также масштабный социально-философский анализ этического и правового регулирования применения ИИ.

4. Значимость для науки и практики выводов и рекомендаций диссертанта.

Теоретическая значимость результатов диссертационного исследования М.Х. Бадмаевой заключается в том, что в ней впервые с позиций социальной философии, ориентированной на достижение высших гуманистических идеалов развития общества, исследованы проблемы разработки, внедрения и применения систем ИИ, сформулировано и уточнено содержание философских принципов безопасного взаимодействия человека и общества с искусственным интеллектом. Это позволит в дальнейшем определить основной вектор развития этих систем, будет способствовать их совершенствованию и эффективному использованию человеком, избегая возможного ущерба для общества и человека.

Исследованный в рамках диссертации и введенный в научный оборот массив документов в виде национальных стратегий, законодательных актов, деклараций, этических руководств, стандартов, резолюций может быть использован для совершенствования и развития этического и правового регулирования в данной сфере на национальном и международном уровне, а также для разработки учебных и учебно-методических курсов по мировоззренческим и методологическим проблемам применения ИИ, крайне востребованных сегодня на всех уровнях профессионального образования, особенно в технических вузах.

Практическое значение данного исследования состоит в выработке ряда практико-ориентированных рекомендаций, которые могут быть успешно использованы в самых различных сферах общественной жизни, в которых актуально внедрение и применение ИИ. Использование результатов исследования в практической сфере будет способствовать повышению результативности и эффективности государственного регулирования ИИ, обеспечению национальной безопасности и процветания государства. Диссертационное исследование может послужить источником необходимой информации для специалистов, работающих в сфере регулирования новых технологий, таких как блокчейн технологии, NBICS-технологии, квантовые вычисления. Кроме того, материалы диссертационного исследования могут быть использованы для формирования профессиональной компетентности специалистов высшей квалификации в сфере ИИ и научных сотрудников, осуществляющих исследования в области разработки и применения систем искусственного интеллекта.

5. Дискуссионные вопросы и замечания

При общей положительной оценке диссертационной работы Бадмаевой М.Х., можно отметить, что некоторые из положений, представленных в диссертации, носят дискуссионный характер.

1. Как отмечается в тексте и в автореферате диссертации, объектом исследования выступают системы слабого, или узкого ИИ (с.14 автореферата диссертации). В этой связи в некоторой степени спорным представляется тезис об экзистенциальной угрозе человечеству со стороны данных систем (с.17 автореферата диссертации). Нам представляется, что проанализированные автором социальные проблемы широкого внедрения в нашу жизнь систем ИИ в качестве составляющей научно-технологического развития и глобальной информатизации общества, такие как, например, некорректное внедрение, нарушение автономии, усиление неравенства, социальная несправедливость, отсутствие ответственности и др., несомненно, требуют скорейшего осмысления и решения, но вряд ли могут быть охарактеризованы как экзистенциальные угрозы, а существующие технологии слабого ИИ, пока во всяком случае, не приводят к пересмотру человеком своего места в мире, искажению его статуса как единственного познающего субъекта, отчуждению от собственной личности. Таким образом, существующие в настоящее время проблемы применения ИИ носят больше социально-этический, чем онтологический и гносеологический характер. Это подтверждает и сам автор, справедливо ограничивая рамки своего исследования социально-философской и этической проблематикой.

2. Автору следовало бы для репрезентативности полученных результатов более подробно изложить данные о том, на основании каких критериев отбирались документы этического и правового характера, представленные в Приложении к диссертации.

3. Анализируя в параграфе 3.2. полученные результаты (в виде выявленных принципов безопасного и эффективного применения ИИ), автор указывает именно шесть принципов применения ИИ. Следовало бы указать, на каком основании был выявлен и отобран именно этот ряд принципов.

Приведенные замечания не меняют общего высокого уровня научного исследования М.Х. Бадмаевой и носят частный характер. Текст четко структурирован, изложен хорошим научным стилем, грамматически выверен. Необходимо отметить научную культуру диссертации, строгость категориального аппарата, научную добросовестность автора. Все выводы четки и обоснованы, подтверждены фактами и статистическими данными. При этом в диссертации нет ортодоксальности, излишне категоричных и непродуманных суждений. Структура диссертации логична, между главами существует органическая связь и преемственность.

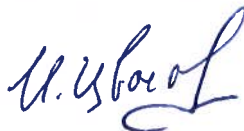
6. Заключение о соответствии диссертации критериям, установленным Положением о присуждении ученых степеней

Диссертационная работа Бадмаевой Майны Харлановны является законченным, научно-квалификационным трудом, раскрывающим заявленную тему, достигающим сформулированные в работе цель и задачи исследования и имеющим существенное значение для области научных интересов соискателя. По теме диссертации опубликовано 8 научных работ, отражающих основные результаты диссертации, из них в изданиях, включенных в Перечень рецензируемых научных журналов, в которых должны быть опубликованы основные научные результаты диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук (Перечень ВАК): 4 статьи. Публикации автора по теме диссертации раскрывают положения, выносимые на защиту. Автореферат диссертации отражает ее содержание.

Диссертационная работа «Социально-философские проблемы и принципы применения систем искусственного интеллекта» соответствует п.п. 9-11, 13, 14 Положения о присуждении ученых степеней, утвержденного постановлением Правительства Российской Федерации от 24.09.2013 г. № 842, а ее автор Бадмаева Майна Харлановна заслуживает присуждения ученой степени кандидата философских наук по специальности 5.7.7. – Социальная и политическая философия.

Официальный оппонент:

доктор философских наук,
профессор кафедры философии
ФГБОУ ВО «Московский авиационный
институт (национальный
исследовательский университет)»



Цвык Ирина Вячеславовна

8 июня 2023 г.

Сержис Цыкис Н.В. завершено.

Федеральное государственное бюджетное
образовательное учреждение
высшего образования
«Московский авиационный институт
(национальный исследовательский университет)»
Адрес: 125993, г. Москва, Волоколамское шоссе, 4.
Телефон: +74991584333,
e-mail: mai@mail.ru
сайт: <https://www.mai.ru>

