

ОТЗЫВ

официального оппонента, кандидата философских наук,
доцента кафедры изобразительного искусства и культурологии
ФГБОУ ВО «Пензенский государственный университет»
Мешковой Людмилы Николаевны
на диссертацию Бадмаевой Маины Харлановны
«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)

Актуальность диссертационного исследования. В настоящее время происходят значительные изменения в обществе и культуре в связи с проникновением цифровых технологий в самые различные области деятельности человека. Одним из направлений цифрового развития является использование систем искусственного интеллекта (ИИ). Широкое внедрение данной технологии в разные сферы производства заставляет по-новому осмыслить проблему «человек и техника», так как системы ИИ обладают такими характеристиками, какие раньше приписывались только самому человеку, например, способность анализировать, обобщать, выводить интеллектуальные решения; свойства обучаться, самостоятельно заниматься поиском информации и принимать на ее основе соответствующие решения (стр. 27). В связи с этим становится важным осмысление социокультурных последствий применения искусственного интеллекта, а также разработка принципов дальнейшего взаимодействия человека и систем ИИ, которые бы учитывали гуманистические идеалы и ценности. Безусловно, данная проблема уже становилась объектом социально-философского и этико-правового анализа. Однако следует признать, что социально-философская рефлексия пока отстает от быстроразвивающихся и всё более широко используемых новых технологий. Кроме того, большинство исследований посвящено отдельным проблемам, возникающим в отношениях человека с ИИ, что не позволяло выработать адекватную и безопасную стратегию взаимодействия общества и интеллектуальных технологий.

В связи с этим представленное диссертационное исследование, направленное на социально-философский анализ проблем применения искусственного интеллекта в современном обществе и выработку содержания базовых принципов взаимодействия человека и систем ИИ для достижения целей гуманистического, безопасного и эффективного развития современного социума, является актуальным и представляет несомненную научную ценность для развития социально-философского знания.

Обоснованность и достоверность результатов диссертационного исследования обеспечивается грамотным использованием системного и деятельностного подходов; принципа историзма; герменевтического метода и общенаучных методов (анализа, синтеза и др.); теорий российских и зарубежных исследователей как методологической основы поиска ответов на поставленные вопросы; анализом и обобщением опыта применения систем искусственного интеллекта в разных областях. Источниковая база рассматриваемого исследования характеризуется широтой и многообразием (253 наименования), включает отечественные и зарубежные работы в области социальной философии, философии техники, аналитической философии, социологии, теории искусственного интеллекта. Кроме того, соискатель обращается к изучению документов, регламентирующих этико-правовые аспекты использования ИИ, созданные и опубликованные государственными структурами, международными и неправительственными организациями, как отечественными, так и зарубежными. Это позволяет автору целенаправленно решать поставленные задачи и убедительно аргументировать собственные выводы.

Научная новизна (стр. 17-18) и положения, выносимые на защиту (стр. 18-21), развернуты и детально обоснованы, адекватно отражают значимые результаты для науки и практики, полученные диссертантом. Выдвинутая автором цель исследования – «выявить и раскрыть содержание философских принципов взаимодействия систем ИИ с человеком для повышения безопасности и снижения рисков их использования в различных сферах

общественной жизни» (стр. 13) – сформировала общую логику научного анализа.

Необходимо отметить логичность построения диссертационной работы, где в трех главах последовательно рассматриваются сущность искусственного интеллекта, его области применения и особенности исследования, этическое и правовое регулирование искусственного интеллекта, выделяются и анализируются социально-философские проблемы применения ИИ в современном обществе. Это явилось основанием выявления и раскрытия смысла философских принципов разработки, внедрения и применения систем искусственного интеллекта.

В качестве достоинств работы следует указать умение работать с понятийно-категориальным аппаратом, четко излагать свои мысли и обозначать свою позицию, выстраивать содержательные и логические связи между частями работы, последовательно обобщать получаемые результаты, что характеризует высокий уровень академической культуры соискателя.

Полученные научные результаты и выводы отвечают поставленным задачам.

Выполненное диссертационное исследование отличается *научной новизной*. В качестве наиболее значимых показателей научной новизны выделим следующие:

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

- обоснована необходимость исследования проблем применения систем ИИ с позиции социальной философии, которое позволяет подойти к их решению с позиций целостности социальной жизни и нацеленности на индивида;

- выделены проблемы, возникающие вследствие применения систем ИИ, которые разделены на две группы. Во-первых, это социальные проблемы, отражающие основные проявления отрицательного влияния ИИ на человека: проблемы причинения вреда, социальной несправедливости и нарушения автономии человека. Во-вторых, это проблемы, раскрывающие причины возникновения нежелательных для человека и общества последствий взаимодействия с ИИ: проблемы непрозрачности, отсутствия ответственности и нарушения конфиденциальности;

- выявлено и раскрыто содержание философских принципов, отражающих условия безопасного и эффективного использования ИИ в различных областях жизни современного общества: принципы прозрачности, конфиденциальности, ответственности, а также принципы социальной справедливости, автономии человека и непричинения вреда. Обосновано основополагающее значение принципа непричинения вреда во всей совокупности указанных философских принципов.

Значимость для науки и практики полученных результатов.
Теоретическая значимость представленного диссертационного исследования состоит в анализе и систематизации основных проблем, связанных с применением ИИ, с позиции социально-философского понимания роли и места человека в мире и выработке на основе этого содержания философских принципов применения систем ИИ. Анализ содержания диссертации позволяет говорить о вкладе М.Х. Бадмаевой в приращении знаний в области социальной философии и теории искусственного интеллекта, а именно в философском осмыслении и анализе социальных последствий, которые обусловлены развитием и распространением применения систем ИИ.

Практическая значимость исследования состоит в том, что его результаты могут стать основой нормотворческой деятельности в области искусственного интеллекта и руководством для дальнейшего этического регулирования процессов взаимодействия человека с ИИ. Также полученные

выводы могут быть использованы в образовательной деятельности при подготовке учебных курсов, где рассматриваются вопросы взаимодействия человека и техники.

Дискуссионные вопросы и замечания. Положительно оценивая диссертацию М.Х. Бадмаевой, следует отметить *ряд замечаний*, которые, на наш взгляд, нацелены на углубление представленного автором исследования.

1. В своей работе диссертант поднимает проблему доверия (стр. 111). Речь здесь идет о доверии к результатам работы системы ИИ со стороны человека. Однако представляется, что проблема доверия имеет и иную сторону. Вполне вероятно, что результаты работы конкретной системы ИИ могут приниматься обычными пользователями без особого сомнения. Быстрота обработки информации, оперативная работа с большим массивом данных могут «заворожить» пользователей. И, одновременно, такая ситуация может породить недоверие к способностям человека, занижение его возможностей, вызвать «комплекс неполноценности» у конкретного пользователя. Таким образом, можно предположить, что ИИ повлияет на самооценку личности, а, соответственно, на ее самореализацию. Возникает вопрос: к какому из выделенных проявлений отрицательного влияния ИИ на человека можно отнести описанную ситуацию?

2. Диссертант выделяет шесть философских принципов разработки, внедрения и применения систем ИИ и отмечает, что необходимость решения задач исследования обусловила разделение этой совокупности на две группы (стр. 140). Однако автор не указывает четко критерии, по которым он выделяет эти группы. Их можно выявить только из контекста исследования.

3. В связи с предшествующим замечанием возникает еще одно. Опираясь, в частности, на задачи исследования (стр. 14), можно предположить, что первая группа включает в себя философские принципы, нацеленные на выявление условий эффективного и безопасного применения систем ИИ в современном обществе, а вторая группа – философские принципы, раскрывающие отдельные стороны негативного влияния ИИ на

жизнь человек и общества. В то же время входящий во вторую группу шестой принцип – принцип непричинения вреда – характеризуется автором как играющий основную роль во всей совокупности представленных принципов, как являющийся системообразующим по отношению к ним (стр. 185). Не является ли внутренним противоречием отнесение системообразующего принципа непричинения вреда к принципам, раскрывающие лишь отдельные стороны негативного влияния ИИ?

Высказанные в данном отзыве замечания не снижают общего высокого качества рассмотренной работы.

Заключение о соответствии диссертации критериям, установленным Положением о порядке присуждения ученых степеней.

Диссертация М. Х. Бадмаевой «Социально-философские проблемы и принципы применения систем искусственного интеллекта» представляет собой самостоятельное, законченное научное исследование, отражающее авторскую точку зрения. В нем были решены задачи, имеющие значение как для социальной философии, так и для осмысления феномена искусственного интеллекта в целом.

Основные положения и результаты работы апробированы и нашли отражение в научных трудах автора. Автореферат и публикации (8 наименований, в том числе 4 – в журналах из перечня ВАК РФ) соответствуют основному содержанию диссертации.

Работа соответствует паспорту специальности 5.7.7. Социальная и политическая философия (философские науки) в части п. 43. Социально-философские проблемы развития информационных технологий и обществ. Перспективы постинформационного общества; п. 44. Социально-философские аспекты развития искусственного интеллекта и цифровизации. Постгуманизм как социально-философская проблема.

Таким образом, анализ диссертации и автореферата позволяют сделать вывод о том, что диссертационное исследование «Социально-философские проблемы и принципы применения систем искусственного интеллекта»

соответствует требованиям п.п. 9-11, 13, 14 «Положения о порядке присуждения ученых степеней», утвержденного постановлением Правительства Российской Федерации от 24.09.2013 г. № 842, предъявляемым к диссертациям на соискание ученой степени кандидата наук, а ее автор Маина Харлановна Бадмаева заслуживает присуждения искомой ученой степени кандидата философских наук по специальности 5.7.7. – Социальная и политическая философия (философские науки).

«09» июня 2023 г.

Официальный оппонент:

кандидат философских наук, доцент,
доцент кафедры изобразительного искусства
и культурологии
ФГБОУ ВО «Пензенский государственный
университет»



Л.Н. Мешкова



Мешковой Л.Н.

Подпись _____ заверяю

Академический секретарь Ученого Совета

О.С. Дорофеева

Федеральное государственное бюджетное
образовательное учреждение
высшего образования
«Пензенский государственный университет»

Адрес: 440026, г. Пенза, ул. Красная, 40

Тел.: 8 (8412) 66-60-01

E-mail: rector@pnzgu.ru

Официальный сайт организации: <https://pnzgu.ru/>