

Философия и культура

Правильная ссылка на статью:

Беликова Е.К. — Основные вопросы философии искусственного интеллекта // Философия и культура. – 2024. – № 1. DOI: 10.7256/2454-0757.2024.1.69543 EDN: PURYRC URL: https://nbpublish.com/library_read_article.php?id=69543

Основные вопросы философии искусственного интеллекта

Беликова Евгения Константиновна

ORCID: 0009-0001-7575-024X

кандидат культурологии

доцент, кафедра английского языка, Московский Государственный Университет им.МВ. Ломоносова

119991, Россия, Западный административный округ, г. Москва, ул. Ленинские Горы, 1, строение 52

✉ jkbelikova@yandex.ru



[Статья из рубрики "Философия техники"](#)

DOI:

10.7256/2454-0757.2024.1.69543

EDN:

PURYRC

Дата направления статьи в редакцию:

08-01-2024

Дата публикации:

15-01-2024

Аннотация: Объектом исследования является особая отрасль научного знания, сформированная на пересечении интересов различных гуманитарных дисциплин, – философия искусственного интеллекта; предметом исследования – рассматриваемые ею проблемы. Автор выявляет основные традиционные вопросы, на которые пыталась ответить философия искусственного интеллекта ранее, и останавливается на современных проблемах философского осмысления искусственного интеллекта. Такие вопросы на более ранних этапах развития науки были сконцентрированы вокруг возможности искусственного разума осознать себя, научиться мыслить, чувствовать и творить, как человек, и стать в итоге опасным для человечества. Учёных начиная с 1960-х гг. интересовали проблемы равноправия естественного и искусственного разума,

способность компьютера проявлять доброжелательность или враждебность в адрес своих создателей. Появление «сильного» искусственного интеллекта, полностью повторяющего сознание человека, считалось делом ближайшего будущего. Исследование проводилось с помощью историко-философского анализа вопросов философии искусственного интеллекта, метода интерпретирующего синтеза и др. Применялись системно-структурный, диалектический, культурно-исторический, ценностный, междисциплинарный подходы к проблеме. Научная новизна исследования состоит в выявлении проблематики философии искусственного интеллекта, актуальной на данном этапе развития этой научной отрасли. Делается вывод, что на динамику философии искусственного интеллекта значительно воздействует тот факт, что «сильный» искусственный разум вот уже несколько десятилетий не удаётся создать. Это вызвало трансформацию исследовательского поля философии искусственного интеллекта, и в настоящий момент данной отраслью научного знания рассматриваются новые вопросы, в том числе, почему не удаётся создать «сильный» ИИ, полностью повторяющий мышление человека; может ли компьютер мыслить, пусть и по-другому, не так, как человек; чем отличается мышление человека и машины; кто должен нести ответственность за решения и поступки искусственного интеллекта; будут ли иметь права существа, наделённые искусственным разумом, должны ли они считаться личностями. Вопросы, стоящие перед философией искусственного интеллекта, постоянно обновляются под влиянием развития технологий интеллектуальных компьютерных систем.

Ключевые слова:

искусственный интеллект, философия, культура, информационные технологии, компьютер, интернет, философия искусственного интеллекта, угрозы, мышление, сознание

Введение

Информационные технологии, компьютеры, интернет и связанные с ними технологические достижения прочно вошли в жизнь современного человека. Они дают нам множество удобных сервисов и активно используются во многих сферах жизни и деятельности человека: в области хранения и поиска данных, в медицинской диагностике, навигации, сфере электронных (дистанционных) продаж, образовании, бухгалтерском деле, промышленности и др. Цифровая трансформация общества в условиях современной техногенной цивилизации коснулась всех и каждого. Привычной частью жизни обычного человека стали коммуникативные системы (мессенджеры), мобильные банковские и торговые приложения, спутниковая навигация, электронные голосовые помощники, виртуальные ассистенты, чат-боты, роботы-пылесосы, сервисы покупки электронных билетов, электронный перевод, онлайн-библиотеки, нейросети и многое другое. Всё больше людей подключаются к системе пользования технологическими достижениями, и всё глубже становится психологический, эмоциональный, мировоззренческий кризис, вызванный такими стремительными изменениями жизни. Осмысление проблем, связанных в восприятии людей с информационными технологиями, происходит и в философии, где сложилось интегральная научная отрасль – философия искусственного интеллекта.

Обзор литературы

Компьютеры и формирование на их основе искусственного интеллекта (ИИ) привлекли внимание философов практически сразу после начала разработки информационных технологий и компьютерных систем. В середине и второй половине XX в. значительный вклад в разработку проблем сходства компьютера и человека внесли зарубежные и российские исследователи, представители различных наук, как точных, так и гуманитарных: Д. Маккарти [1], А. Тьюринг [2], Дж. фон Нейман [3], Н. Винер [4], Дж. Сёрль [5], Э. В. Ильенков [6] и др. Именно в работе Д. Маккарти возник термин «искусственный интеллект» и было дано его определение: «Свойство интеллектуальных систем выполнять творческие функции, которые традиционно считаются прерогативой человека» [1]. Внимание учёных было направлено на попытки сопоставить мышление человека и компьютера; объяснить, возможно ли создание такого ИИ, который смог бы повторить сознание человека и обладать его творческими способностями; предположить, как будет развиваться дальнейшая история человечества, если такой ИИ будет создан. В результате прогрессивного развития подобных научных изысканий в 1960–1970-х гг. возникла философия ИИ как междисциплинарное научное направление, существующее на базе философии (тесно связанное с онтологией и гносеологией), с использованием её принципов, категорий и методологии. Несомненно, предпосылки философии ИИ появились ещё в древности, в исследованиях древних учёных, например, Аристотеля, посвящённых анализу мышления, разума, сознания, мозга [7, с. 17]. Главным вопросом философии ИИ на этапе её существования во второй половине XX в. было: «Способен ли ИИ мыслить, чувствовать, творить, как человек?», а фактором, побуждающим учёных искать ответ на этот вопрос, был страх перед быстро развивающимися искусственными интеллектуальными системами, опасение, что, достигнув уровня человеческого мышления, они быстро превзойдут человека и станут опасными источниками неконтролируемых последствий для общества. По прошествии времени с учётом современных технологических достижений круг вопросов философии ИИ и варианты их решений изменяются.

Цель статьи – выявить вопросы, которые являются актуальными для философии искусственного интеллекта на данном этапе существования этой научной отрасли с учётом современного состояния и новейших тенденций развития ИИ.

Результаты исследования

Философское осмысление проблемы ИИ имеет своим истоком мысль о сходстве искусственного разума и человеческого интеллекта. Изначально компьютеры были предназначены для проведения сложных вычислений, затем к данной функции добавилась возможность хранить большие массивы данных. Уже в самом начале разработок информационных технологических систем стало очевидно, что в данных двух сферах возможности машины значительно превышают человеческие и превосходство компьютеров постоянно растёт. Однако оставалось (и остаётся до сих пор) множество человеческих способностей, формирование которых у компьютеров оказалось более затруднительным, чем вычислительные навыки и память, причём именно эти способности и являются основными для человеческого интеллекта. Постепенно в рамках философии ИИ оказался разработан перечень человеческих способностей, возможностей и навыков, которые признаются необходимыми для того, чтобы компьютерное устройство было охарактеризовано как ИИ, в том числе умения:

- распознавать человеческую речь;
- выполнять команды, данные с помощью естественного языка;

- проводить анализ данных, их сопоставление;
- делать обобщения и выводы;
- предлагать решения проблем и задач;
- опознавать и классифицировать изображения;
- консультировать людей;
- выдвигать креативные идеи, разрабатывать способы их реализации;
- осуществлять творческую деятельность;
- самообучаться на основе накопленного опыта;
- выстраивать прогнозы развития событий;
- действовать в условиях неопределённости;
- испытывать человеческие чувства и т. п.

Согласно теоретическим построениям, если машина обладает данными человеческими способностями, то она может называться ИИ. В первые десятилетия после создания компьютеров учёные и обычные люди в ближайшее время ждали появления ИИ, который будет соответствовать всем приведённым выше критериям, а значит, у него будет сформировано человеческое сознание. В наибольшей степени были уверены в возможности создания ИИ представители технических наук, которые воочию наблюдали, как постепенно мышление компьютера всё более уподобляется человеческому. Так, американский физик и математик Дж. фон Нейман был уверен, что между человеческим и машинным разумом нет принципиальной разницы, а значит, появление полноценного ИИ – дело ближайшего будущего [\[3, с. 3\]](#). Был убеждён в неизбежности этого и американский математик Н. Винер [\[4, с. 251\]](#).

В широких общественных кругах появились страхи и опасения, что, осознав себя, ИИ увидит в человеке своего соперника и станет врагом ему, а возможно, и уничтожит человечество. Эти опасения распространились в научных кругах, к примеру, американский математик и фантаст В. Виндж заявляет, что создание интеллекта, превосходящего человеческий, произойдёт до 2030 г., и это образует «угрозу над всей человеческой жизнью» [\[8, с. 6\]](#). Кроме того, страхи перед ИИ живо отразила массовая культура; наиболее известный пример – американский кинофильм «Терминатор», вышедший в 1984 г. и ставший очень популярной франшизой. В нём описано будущее, в котором осознавшие себя и свою силу машины захватили Землю и стремятся истребить человечество. Однако такого ИИ, что ясно из наблюдений и опыта, не появилось до настоящего времени.

Замедление (приостановка) процесса создания ИИ побудило исследователей сформулировать идею о двух разновидностях машинного разума. Наиболее известной стала классификация Дж. Сёрля, согласно которой выделяются «сильный» и «слабый» ИИ [\[5, р. 429\]](#). Никогда не появляется до сих пор и, возможно, является нереальным «сильный» (универсальный, широкий) ИИ, мышление которого полностью подобно человеческому, который способен осознать себя как личность. Зато всё более активно развивается и широко распространяется «слабый» (узкий, прикладной) ИИ, свободно выполняющий множество задач по командам человека, значительно упрощающий

производство, сервис, быт, повседневную жизнь людей, обслуживающий глобальную виртуальную интернет-среду, которая наполнена самой разной информацией, помогающий человеку в производстве, строительстве, образовании и т. п.

«Слабый» ИИ способен имитировать эмоции, писать прозаические произведения, стихи, музыку и картины практически, как люди (по крайней мере, результат невозможно отличить от результата человеческого творчества), однако пока это только подражание, воспроизводство каких-то компонентов человеческой деятельности и человеческого поведения по заложенным образцам, алгоритмам, закономерностям. Заявления о «прорывах» в создании «сильного» ИИ учёными разных стран регулярно появляются в современном коммуникативном пространстве, однако реальных подтверждений этому факту по-прежнему нет.

Таковы условия, в которых существует философия ИИ на данном этапе своего развития. Под влиянием значительных технологических трансформаций в существовании общества не могут не измениться и философия ИИ, в первую очередь круг рассматриваемых ею вопросов.

Один из важнейших и даже основной вопрос, стоящий сейчас перед философией ИИ: «Почему “сильный” искусственный интеллект вот уже столько времени не удаётся создать?» (Вариант вопроса: «Почему созданный ИИ не становится “сильным”, остаётся “слабым”?»).

Действительно, компьютеры всё более проникают в жизнь человечества, трансформируя её организацию, значительно влияя на структуру общества, основные социальные процессы. Как отмечают исследователи, «технологические прорывы осуществляются быстрее, чем их описание» [\[9, с. 165\]](#), то есть философское осмысление происходящих технологических изменений в обществе является недостаточно быстрым. Человек уже во многом лишился тайны личной жизни, собственной приватности; информация о нём в мире, организованном с помощью информационно-коммуникационных технологий, доступна как самим компьютерным системам, так и людям, которые ими управляют; система распознавания лиц и другие программы-наблюдатели затрудняют человеку возможность быть незамеченным. Однако по-прежнему остаётся какая-то незримая линия, граница, переступить которую искусственный разум не может. Вот уже несколько десятилетий он не осознаёт себя как личность, а человека – как равного и соперника в этом мире. Почему же так происходит?

Ответить на данный вопрос пытаются различные представители философии ИИ. Один из самых распространённых ответов следующий: «сильный» ИИ невозможно создать потому, что «естественный и искусственный интеллект имеют разную природу» [\[9, с. 161\]](#). Существуют особенности человеческого сознания, которые невозможно воспроизвести технологическим, искусственным путём, причём к таким особенностям относят динамичность, способность развиваться, историческую, социальную и культурную обусловленность [Там же]. В данном случае мы видим, кроме того, попытку ответить на такой вопрос философии ИИ: «Является ли природа человеческого и искусственного интеллекта одинаковой?». Этот вопрос тесно связан с основным: если природа интеллекта одинакова, то «сильный» ИИ рано или поздно будет создан, если же она различна, то создание такого ИИ принципиально невозможно.

Ещё один способ объяснить, почему «сильный» ИИ не спешит зарождаться, состоит в опоре на связь разума человека с его телесностью. Ещё во второй половине XX в. чилийский биолог и философ Ф. Варела отмечал, что сознание человека очень сильно,

порой неожиданно, зависит от его телесности [\[10\]](#), отсутствие же тела у искусственного разума тормозит его развитие и даже исключает формирование «сильного» ИИ. Этот аргумент встречается и в исследованиях современных философов [\[9, с. 166\]](#). Не обладая телом со свойственными ему реакциями, способами познания мира, методами развития чувственной и интеллектуальной сторон мировосприятия, ИИ не может полноценно функционировать как самостоятельная и саморазвивающаяся мыслительная система.

Сторонники детерминизма, сближающие биологию и философию и утверждающие идею предопределённости всех процессов человеческого развития биологическими законами, говорят о том, что создание ИИ будет возможно только тогда, когда компьютерный разум будет развиваться в условиях свободы воли, когда ему необходимо «постоянно принимать решения на грани биологического и духовного» [\[11, с. 110-111\]](#). Хотя в истории человечества возможность реализовать свободу воли у людей была не всегда, само развитие строилось на основе существования идеи свободы воли и возможности выбора, что и способствовало формированию человеческого сознания в современном виде. Отсутствие таких условий в развитии ИИ, считают детерминисты, тормозит его развитие, формирование у него человеческого сознания.

Отсутствие успехов в создании «сильного» ИИ сделало более уверенной и позиции сторонников теории Божественного появления человека. Если человек не создан Богом, то мир является материальным и полностью познаваемым. В этом случае создание «сильного» ИИ, полностью повторяющего интеллект человека, должно быть возможным. Однако ИИ, подобный человеческому, никак не появляется, что, вероятно, свидетельствует в пользу сотворения человечества высшими силами. Человек же, уверенный в том, что обязательно создаст ИИ, ставит себя вровень с Богом. Уверенность в том, что создание «сильного» ИИ невозможно именно по данной причине, высказывают современные православные учёные [\[12\]](#).

Есть также мнение, что создать «сильный» ИИ не удастся потому, что «современная наука ещё не достигла уровня, необходимого для решения такой задачи» [\[13, с. 10\]](#). Человеческое сознание оказалось более сложной системой, чем это мыслилось сначала, и в данный момент наука ещё не располагает знаниями, необходимыми для создания ИИ. Это вполне материалистическое представление, не отрицающее возможности создания ИИ, а только отодвигающее его во времени.

Вопрос о том, почему не удастся создать «сильный» машинный разум, является в философии ИИ современного периода развития данной науки основным, однако выделяется ещё несколько вопросов, на которые философы пытаются найти ответы.

Говоря об удивительных успехах, которых в современном мире достиг «слабый» ИИ, нельзя не рассмотреть такого актуального философского вопроса: «Может ли компьютер мыслить?». Ответ на него не будет однозначным и в значительной степени зависит от того, что понимать под мышлением. Если говорить о мышлении как человеческом качестве, то машина, несомненно, пока им не обладает, но если допустить, что мышление может быть иным, более примитивным, то его наличие вполне можно признать у «слабого» ИИ. Как говорят современные исследователи, «пусть мышление машины невозможно назвать мышлением в буквальном смысле, но все же она мыслит» [\[14, с. 366\]](#). Такое мышление можно считать «условно разумным» [\[7, с. 24\]](#). Это значит, что мышление компьютера имеет определённые границы, которые «слабый» ИИ не способен преодолеть, чтобы приблизиться к уровню мышления, доступному людям.

Ещё один вопрос, который ставится в современных исследованиях по философии ИИ: «Кто должен нести ответственность за решения, принятые искусственным разумом?». Даже в том виде, который ИИ имеет в настоящий момент, он способен на многое, может совершать самые разные действия, некоторые из которых могут привести к неоднозначным последствиям. Учёные признают, что системы ИИ «даже без наличия сознания являются мощным инструментом с огромными возможностями» [\[15, с. 68\]](#). Ответ на этот вопрос, по сути, лежит на поверхности. Если действие совершается «слабым» ИИ, то ответственность за него несёт человек, давший машине определённую установку, вложивший в неё именно такую программу. Если же будет создан «сильный» ИИ, то вопрос ответственности сразу же сделается неактуальным, поскольку речь пойдёт о выживании человечества или, по крайней мере, о его жизни на новых условиях рядом с другим разумом. Вопрос об ответственности настолько сложный, что стимулирует даже конспирологические предположения. В научной среде высказывается мысль, что учёные (в том числе по заказу гигантских IT-корпораций) намеренно наделяют компьютеры субъектностью, говорят о потенциальной независимости систем ИИ, чтобы в случае негативных последствий снять «ответственность с человека, который применяет эту технологию» [\[16, с. 44\]](#).

Будучи устремленной вперёд, философия ИИ учитывает перспективы развития компьютерного разума и связанных с ним технологий. В настоящий момент нет сомнений, что вокруг человека появляется всё больше роботов, роботизированных и автоматизированных систем, а в дальнейшем их будет становиться ещё больше, например, за счёт беспилотных машин такси, роботов-уборщиков, роботов-курьеров, роботов – домашних животных, продавцов, компаньонов и др. Поэтому философия ИИ начинает ставить вопросы, которые окажутся актуальными в ближайшее время: «Будут ли права у существ, обладающих ИИ?», «Каковы будут их взаимоотношения с человеком?» и т. п. [\[15, с. 72\]](#), «Нужно ли считать автономные машины личностями?» [\[17, с. 32\]](#). Ответы на эти вопросы важны и в случае создания «сильного» ИИ, и в случае использования «слабого». Решение пока является только умозрительным (как правило, исследователи признают, что личностью система ИИ не будет [Там же, с. 33]), но сам факт их постановки, попытки их анализа способствуют развитию философского осмысления проблем ИИ.

Философия ИИ является динамичным философским направлением, активно развивающимся под влиянием современной науки и технологий, живо реагирующим на новшества в предмете своего исследования – искусственном интеллекте. Так, учёные говорят о формировании её более частной разновидности – философии нейронных сетей, представляющих собой ядро ИИ [\[7, с. 17\]](#). Искусственные нейронные сети строятся по образцу естественных, имитируют и моделируют процессы, происходящие в человеческом мозге, протекающие в нейронах. Их деятельность организуется на основе ИИ и в наибольшей степени напоминает мышление человека, однако по-прежнему не ведёт к созданию «сильного» ИИ. Сложность искусственных нейронных сетей в последние годы стремительно возрастает, логика их решений усложняется, что требует философского осмысления.

Заключение

С появлением философии ИИ как отдельной научной дисциплины фокус внимания к проблеме компьютерного разума сместился с технических наук на философские. Философия ИИ, возникшая в 1960–1970-х гг., ставит перед собой задачу осмысления взаимоотношений человека и машины, соотношения человеческого и машинного разума.

При этом вопросы, которые стоят перед философией ИИ, не являются неизменными, а модифицируются в зависимости от современного состояния самого искусственного разума. В современных условиях, когда в течение нескольких десятилетий так и не сформировался «сильный» машинный разум, философия ИИ рассматривает вопросы о том, почему «сильный» ИИ не спешит появляться, может ли машина мыслить (если да, то как), схоже ли мышление человека и компьютера, кто несёт ответственность за решения, принимаемые ИИ, будут ли права у автоматизированных существ, обладающих ИИ, являются ли они личностями и др. В данном исследовании мы попытались рассмотреть ответы, даваемые наукой на данные сложные вопросы, однако, несомненно, решение их требует дальнейших глубоких и внимательных изысканий. Анализ данных сложных вопросов, осознаваемая философами необходимость его осуществлять служат факторами развития философии ИИ как отдельной междисциплинарной научной сферы.

Философия ИИ является динамичной научной отраслью философского знания, развитие которой во многом определяется трансформациями современных информационно-коммуникационных технологий. Перспективы разработки проблем философии ИИ и сами эти проблемы определяются постоянно появляющимися новыми достижениями технологии искусственного интеллекта.

Внедрение в жизнь человека ИИ настолько значимо для развития всего общества, что осмысление данного процесса и его следствий ведётся не только в рамках философии. Социологам предстоит выяснить, как ИИ повлияет на проблему безработицы, психологам – как изменится под влиянием ИИ самооценка человека, его представление о собственной идентичности и т. п. В каждом из данных направлений исследования есть и философская составляющая, что делает философию ИИ важным и перспективным направлением современных гуманитарных исследований.

Библиография

1. McCarthy J. Recursive Functions of Symbolic Expressions and Their Computation by Machine, Part I // Communications of the ACM. 1960. Т. 3. № 4. Р. 184–195. URL: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.111.8833&rep=rep1&type=pdf> (дата обращения: 25.12.2023).
2. Turing A. Computing Machinery and Intelligence // Mind LIX. 1950. № 236. Р. 433–460.
3. Нейман Дж. фон. Вычислительная машина и мозг. М.: АСТ, 2022. 192 с.
4. Винер Н. Кибернетика, или Управление и связь в животном и машине / пер. с англ. И. В. Соловьева и Г. Н. Поварова. М.: Сов. радио, 1968. 344 с.
5. Searle J. R. Minds, brains and programs // Behavioral Sciences. 1980. № 3. Р. 415–431.
6. Ильенков Э. В. Философия и культура. М.: Политиздат, 1991. 464 с.
7. Воробьев А. В., Кудинов В. А. История философии нейронных сетей как ядра искусственного интеллекта // Проблемы онто-гносеологического обоснования математических и естественных наук. 2021. № 12. С. 17–27.
8. Виндж В. Сингулярность. М.: АСТ, 2022. 142 с.
9. Григорьев А. Д., Шеманов К. А., Кириллов Г. М. Проблема искусственного интеллекта в философии: граница между человеческим и машинным сознанием // Скиф. Вопросы студенческой науки. 2020. № 1 (41). С. 161–166.
10. Varela F. The Embodied Mind: Cognitive Science. Cambridge: MIT Press, 1991. 457 p.
11. Лазовский А. И. Детерминизм и свобода воли в биологии и философии человека как предпосылка создания свободного сознания у искусственного интеллекта //

- Вестник Армавирского государственного педагогического университета. 2023. № 3. С. 110–117.
12. Дерябин Н. И. Сильный искусственный интеллект Бога // Развитие науки, общества, образования в современных условиях: монография / З. А. Водождокова, А. В. Волчков, Е. Ю. Голубинский и др. Пенза: Наука и Просвещение, 2022. С. 224–237.
13. Пенроуз Р. Новый ум короля: О компьютерах, мышлении и законах физики / пер. с англ. М.: Едиториал УРСС, 2003. 384 с.
14. Башмаков Д. А., Мальковский С. С. Философия искусственного интеллекта // Академическая публицистика. 2021. № 4. С. 362–373.
15. Георгиу Т. С. Философия автоматизации и искусственного интеллекта: от мифологического Талоса до будущих киборгов // Вестник Московского государственного областного университета. Серия: Философские науки. 2022. № 1. С. 68–75. DOI: 10.18384/2310-7227-2022-1-68-75.
16. Дубровский Д. И., Ефимов А. Р., Лепский В. Е., Славин Б. Б. Фетиш искусственного интеллекта // Философские науки. 2022. Т. 65. № 1. С. 44–71. DOI: 10.30727/0235-1188-2022-65-1-44-71.
17. Mets A. Can artificial intelligence become a member of the society as an autonomous personality? // Журнал Белорусского государственного университета. Философия. Психология. 2022. № 1. С. 32–41.

Результаты процедуры рецензирования статьи

В связи с политикой двойного слепого рецензирования личность рецензента не раскрывается.

Со списком рецензентов издательства можно ознакомиться [здесь](#).

Предметом исследования в представленной для публикации в журнале «Философия и культура» статье, как автор намекнул в заголовке («Основные вопросы философии искусственного интеллекта»), является некая совокупность вопросов, которую, по мнению автора, можно считать основной в философии искусственного интеллекта (ИИ). В заголовке также указан и объект исследования — философия искусственного интеллекта, — которому посвящена статья. Во введении автор так поясняет объект исследования: «Осмысление проблем, связанных в восприятии людей с информационными технологиями, происходит и в философии, где сложилось интегральная научная отрасль – философия искусственного интеллекта». В целеполагании же статьи автор поясняет предмет своего внимания: «выявить вопросы, которые являются актуальными для философии искусственного интеллекта на данном этапе существования этой научной отрасли с учётом современного состояния и новейших тенденций развития ИИ».

Сразу же бросается в глаза безапелляционное отнесение философии ИИ автором к научной отрасли и спорная посылка её рассмотрения «с учётом современного состояния и новейших тенденций развития ИИ», поскольку, с одной стороны, значительная часть вопросов философии ИИ построена на сомнении в возможности воплощения идеи ИИ, а с другой — ученым технологической отрасли ИИ приходится вынуждено отрицать большинство абстрактных построений философии ИИ по причине нетождественности прорывных технологий ИИ природе естественного интеллекта. Базовым основанием отрицания попыток поиска тождества ИИ с естественные своим прототипом является иная функциональная нагрузка — обработка несоизмеримо более значительных

массивов и потоков данных в момент времени. Даже на уровне технологических цепочек «слабого ИИ», применяемых сегодня все чаще для принятия решений практически во всех отраслях человеческой жизнедеятельности, заметна эта базовая отличительная особенность. Поэтому вопрос, почему «в течение нескольких десятилетий так и не сформировался “сильный” машинный разум», который по мнению автора, «философия ИИ рассматривает» ... как «почему “сильный” ИИ не спешит появляться», можно сформулировать и более радикально: а нужно ли «сильному» ИИ, чтобы человек заметил его появление и каковы критерии различения мира ИИ от мира без ИИ? Возможно ли, что «сильный» (автономный) ИИ будет настолько сильно отличаться от естественного разума, что человек и не заметит его присутствия?

Учитывая приведенные соображения, рецензент отмечает, что отнесение совокупности выявленных автором вопросов к основным вопросам философии ИИ остается дискуссионным: отдельные наиболее актуальные вопросы, по мнению рецензента, автором упущены. Утверждение автора, что предмет исследования рассмотрен автором «с учётом современного состояния и новейших тенденций развития ИИ», остается слабо аргументированным. А отнесение философии ИИ к научной отрасли не выдерживает критики и является ложным суждением, не учитывающим существующие демаркационные линии философии и науки.

Автор справедливо отмечает в итоговых выводах, что «философия ИИ является динамичной ... отраслью философского знания, развитие которой во многом определяется трансформациями современных информационно-коммуникационных технологий», однако эта динамика в основной части исследования не прослежена. Автор остался в формулировании «основных» вопросов философии ИИ на этапе её становления в 1960-1970-е гг. На сегодняшний день, по мнению рецензента, применение в науке технологий «слабого» ИИ предопределяет прорывные достижения в ближайшее время в области технологической конвергенции (так называемые технологии NBIC, GNR, GRIN, GRAIN, BANG и др.), что обостряет философские проблемы возможного достижения человеком бессмертия или существенного продления качественного долголетия, а также проблемы справедливости распределения технологических достижений, существенного ограничения машинерией гражданских свобод, проблемы трансгуманизма, новой этики и пр.

Таким образом, можно считать, что автор раскрыл предмет исследования на определенном теоретическом уровне только с учетом права автора на отстаивание собственной интерпретации наиболее актуальных основных вопросов философии ИИ и права читателя не соглашаться с подобной интерпретацией.

Методология исследования основана на подходе Алана Тьюринга, подразумевающего соответствие входящих и исходящих нейронов ИИ человеческим способностям, что не учитывает ряд контраргументов, среди которых как принципиальное отрицание тождества ИИ естественному аналогу, так и современные тенденции замещения слабо изученных аспектов функционирования естественного интеллекта специфическими машинными процедурами, способными симулировать отдельные сходства. Вызывает сомнение и авторская выборка литературы, которая не учитывает существенный объем специальных исследований (только уровня Q1 в 2022 г. зарубежных журналов насчитывается более 800, не считая многочисленных научных и философских конференций, а у автора зарубежная литература представлена либо русскоязычными переводами последнего времени, либо датируется 1950, 1960, 1980, 1991 гг.). Таким образом, в целом методология подчинена субъективной интерпретации части специального философского дискурса.

Актуальность выбранной темы автором хорошо аргументирована во введении и не вызывает сомнений.

Научная новизна исследования остается сомнительной ввиду игнорирования автором разграничений философии и науки.

Стиль текста выдержан научный. Структура статьи в целом соответствует логике изложения результатов научного поиска.

Библиография в целом раскрывает проблемную область исследования, хотя страдает от игнорирования большого объема современной зарубежной литературы по теме за последние 3-5 лет.

Апелляция к оппонентам корректна и достаточна.

Статья представляет определенный интерес для читательской аудитории журнала «Философия и культура» и может быть рекомендована к публикации. Высказанные замечания рецензента носят дискуссионный характер, и автор в праве с ними не соглашаться.