

УДК: 165.1 + 004.8

DOI: 10.15372/PS20250516

EDN: JDJFTO

О.И. Елхова

**«ЧЕЛОВЕК В КОНТУРЕ»
И ПРЕДЕЛЫ РАЦИОНАЛЬНОСТИ
ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**

В статье рассматриваются философские основания и ограничения рациональности искусственного интеллекта в контексте принятия решений. Автор анализирует различие между эпистемической и практической рациональностью, акцентируя внимание на последней как основе работы рациональных агентов. Центральное место занимает концепция ограниченной рациональности, согласно которой принятие решений осуществляется в условиях неполной информации, когнитивных ограничений и ограниченных вычислительных ресурсов. Обосновано, что идеальная рациональность недостижима, а наиболее адекватной моделью для ИИ является ограниченная рациональность. Рассмотрены четыре типа рациональности: *идеальная*, *вычислительная*, *ограниченная* и *ограниченная оптимальность*, с выводом о практической применимости именно *ограниченной рациональности* в разработке интеллектуальных систем. Сделан вывод о практической применимости ограниченного подхода при разработке интеллектуальных систем. Автор отмечает, что необходимо учитывать указанные ограничения при создании адаптивных алгоритмов, а также значимость участия человека в процессе принятия решений для повышения надёжности результатов. Модель «человека в контуре» трактуется не просто как техническая форма взаимодействия, а как выражение ситуационной рациональности, предполагающей учёт контекста, моральных последствий и уникальности каждой конкретной ситуации. Присутствие человека придаёт алгоритмическому мышлению ценностный и интерпретативный характер, восстанавливая связь между рациональностью и практической мудростью (*phronesis*). В условиях, когда алгоритмы

ограничены ресурсами и подвержены ошибкам, именно человек способен выявлять контекстуальные нюансы и осуществлять осмысленную корректировку решений. Таким образом, включение человека в контур является важнейшим фактором, обеспечивающим повышение надежности решений, принимаемых искусственным интеллектом.

Ключевые слова: человек в контуре, ограниченная рациональность, искусственный интеллект, практическая рациональность, рациональный агент, принятие решений

O.I. Elkhova

«HUMAN IN THE LOOP» AND THE LIMITS OF ARTIFICIAL INTELLIGENCE'S RATIONALITY

The article examines the philosophical foundations and limits of artificial intelligence rationality within the context of decision-making. The author analyzes the distinction between epistemic and practical rationality, emphasizing the latter as the fundamental basis for rational agent functioning. Central to the discussion is the concept of bounded rationality, according to which decision-making occurs under conditions of incomplete information, cognitive constraints, and limited computational resources. The article argues that ideal rationality is unattainable, making bounded rationality the most appropriate model for artificial intelligence. Four types of rationality are explored: ideal, computational, bounded rationality, and bounded optimality. The study concludes that bounded rationality is practically most applicable in developing intelligent systems. Furthermore, it highlights the necessity of considering these limitations when designing adaptive algorithms and underscores the significance of human involvement in the decision-making process to enhance reliability. The «*human-in-the-loop*» model is interpreted not merely as a technical mode of interaction, but as an embodiment of situational rationality, taking into account context, moral implications, and the uniqueness of each specific situation. Human participation adds a value-based and interpretative dimension to algorithmic reasoning, restoring the connection between rationality and phronesis. Given that algorithms are limited by resources and prone to errors, it is precisely the human who is capable of identifying contextual nuances and providing meaningful corrections. Thus, human inclusion in the loop emerges as a critical factor in enhancing the reliability of artificial intelligence-driven decisions.

Keywords: human-in-the-loop, bounded rationality, artificial intelligence, practical rationality, rational agent, decision-making

Введение

Концепция рациональности в философии восходит к Аристотелю, который выделял два ее вида: эпистемическую и практическую. Первая связана с обоснованностью убеждений, вторая – с действиями и принятием решений. Такое разграничение отражает различие между процессом рассуждения и фактическим поведением. В эпистемическом смысле рациональность предполагает наличие аргументированного и надежно сформированного убеждения. Однако даже противоречивые суждения могут сохранять рациональный характер, если они обоснованы с точки зрения используемых методов познания.

Аристотель подчеркивает, что процесс рассуждения и само поведение представляют собой разные феномены. В настоящее время большее внимание уделяется рациональному поведению, поскольку оно связано с практическими результатами, для его описания используется понятие «рационального агента» – субъекта, принимающего решения, направленные на достижение оптимального результата, в том числе в условиях неопределенности. Практическая рациональность существенно отличается от эпистемической рациональности, которая сосредоточена на соответствии рассуждений логическим принципам. В отличие от эпистемической, она охватывает не только логические аспекты, но и мотивацию, желания, учитывает реальные ограничения, направлена на принятие решений и выбор наилучшего варианта действий в конкретных условиях. Именно эта функция закладывается в основу работы искусственного интеллекта, который должен анализировать данные, прогнозировать возможные последствия и определять наиболее эффективные решения. Такой процесс во многом напоминает человеческое практическое мышление, ориентированное на достижение целей, а не только на обоснование убеждений. Понятие «рационального агента» стало ключевым в дискурсе об искусственном интеллекте [1, с. 11–16]. В контексте **искусственного интеллекта** под рациональным агентом полагается система, способная анализировать данные, прогнозировать последствия действий и выбирать оптимальное решение с учетом заданных параметров в условиях неопределенности. В настоящей работе целью исследования является выявление философских ос-

нований и пределов рациональности искусственного интеллекта, поскольку в современных условиях все более актуальным становится детальный анализ его ограничений при обработке информации и принятии решений. Активное внедрение интеллектуальных систем в критически важные сферы жизнедеятельности человека требует пересмотра границ их рациональности и факторов, влияющих на точность и обоснованность принимаемых решений.

Теоретический анализ

Несмотря на то, что искусственный интеллект изначально создавался для моделирования человеческого интеллекта, стоит признать, что решение человека не всегда может быть рациональным и приводить к оптимальным результатам, зачастую оказывается предвзятым. В этой связи стоит вспомнить концепцию *ограниченной рациональности* Г. Саймона, объясняющую процесс принятия решений через призму когнитивных ограничений индивидов и неполноты доступной информации. Данные взгляды существенно отличаются от традиционных представлений, когда человек рассматривается как полностью рациональный агент, способный всегда выбирать наилучший вариант, основываясь на объективном анализе. Классическая теория рациональности, словно не замечая своих скрытых предпосылок, исходит из убеждения, что индивиды обладают исчерпывающим знанием всех возможных альтернатив и их последствий. В свою очередь, Г. Саймон подчеркивает, что в реальных условиях люди редко стремятся к абсолютному оптимуму. Вместо этого они руководствуются принципом удовлетворительности, предпочитая решения, которые кажутся достаточно хорошими, но не обязательно являются наилучшими. В этом контексте рациональность играет ключевую роль, объясняя механизмы принятия решений. Когнитивные ограничения человека Г. Саймон образно сравнивал с одним лезвием ножниц, а структуру окружения – с другим лезвием: разум «режет» эффективно только благодаря взаимодействию ограниченных способностей с подсказками среды [8, с. 198–199]. Теория *ограниченной рациональности* ставит под сомнение пред-

ставление о том, что человек способен принимать решения, опираясь исключительно на объективные данные, без воздействия эмоций и когнитивных искажений, подчеркивая, что рациональность определяется не только итоговым результатом, но и самим процессом выбора. В действительности принятие решений включает анализ доступной информации, учет временных рамок и оценку значимости возможных альтернатив. Когнитивные барьеры, нехватка знаний и влияние внешней среды также препятствуют достижению идеального результата. Актуальность данных положений подтверждается исследованиями, посвящёнными вызовам цифровой эпохи, которые показывают, что развитие современных информационно-коммуникационных технологий, в том числе искусственного интеллекта, порождает принципиально новые вызовы, ставящие под вопрос устоявшиеся представления о рациональности и требующие их фундаментального переосмысления [3, с. 27–30].

Проблема рациональности искусственного интеллекта неразрывно связана с особенностями обработки информации. Такие аспекты проявляются не только в когнитивных ограничениях алгоритмов, но и в специфике взаимодействия человека с цифровой средой, включая виртуальную реальность. В данной среде структурные метрики феноменологического опыта оказывают влияние на субъективные механизмы принятия решений. Анализ феноменологии виртуального опыта показывает, что процесс восприятия цифрового пространства основывается на принципах, схожих с концепцией ограниченной рациональности. В ходе взаимодействия с виртуальной средой человек синтезирует сенсорные данные из обоих миров: физического и цифрового, что способствует формированию адаптивных когнитивных стратегий. Концепция полевой интерференции реального и виртуального в исследовании феноменологического опыта демонстрирует, что рациональность в цифровом пространстве нельзя свести к простому алгоритму оптимизации [2, с. 1003–1005]. Напротив, рациональность формируется под влиянием множества взаимосвязанных факторов: когнитивных, сенсорных и социальных, требующих комплексного подхода к изучению.

В ситуациях, когда количество возможных альтернатив слишком велико для исчерпывающего анализа, индивиды часто прибе-

гают к упрощенным эвристикам, что неизбежно приводит к отклонению от оптимального выбора. Кроме того, субъективные факторы, такие как личные ценности, цели и предпочтения, оказывают значительное влияние на процесс принятия решений, порой искажая объективную оценку возможных вариантов. Ошибки в суждениях могут возникать по следующим причинам: человек не обладает безграничными когнитивными ресурсами и может не знать о более выгодных альтернативах либо, напротив, ошибочно считать невыгодный вариант предпочтительным, т.е. в условиях неопределенности люди удовлетворяются с их точки зрения наиболее подходящим вариантом, а не максимизируют результат. В результате принятые решения не всегда соответствуют объективному максимуму.

Концепция Г. Саймона оказала глубокое влияние на ИИ, было выявлено, что и машины сталкиваются с теми же принципиальными барьерами, что и люди. Теория ограниченной рациональности остается актуальной и в современных исследованиях, особенно при разработке систем с ИИ. Можно выделить два подхода к моделированию рациональности в ИИ. Первый, «снизу-вверх», использует данные и входные сигналы от человека, что может вносить предвзятость. Искажения происходят из-за того, что данные, на которых обучаются модели, несут в себе ложные идеи и ограниченные представления, отражающие человеческие предрассудки. Второй, «сверху-вниз», основан на логике и формальной рациональности, но человеческие предубеждения могут быть заложены в самих правилах модели. Например, в системах машинного перевода могут быть использованы определенные языковые паттерны, которые искажают результаты. Современные исследования активно развивают гибридные подходы, которые пытаются объединить преимущества обеих моделей, сочетая данные и логику, чтобы минимизировать предвзятость и повысить рациональность решений.

В дискуссиях о возможностях искусственного интеллекта выделяют четыре типа рациональности: *идеальную*, *вычислительную*, а также концепции *ограниченной рациональности* и *ограниченной оптимальности* [7, с 36–38, с. 58–59]. *Идеальная рациональность* подразумевает принятие решения на основе полного знания всех возможных альтернатив и их последствий.

Только в реальных условиях такая рациональность остается недостижимой из-за неполноты информации и сложности решаемых задач. Кроме того, некоторые задачи требуют учета огромного количества переменных, что делает их расчет невозможным даже для самых мощных современных систем. В связи с этим более применимой концепцией становится *вычислительная рациональность*, которая учитывает доступные ресурсы и позволяет находить решения, приемлемые в рамках вычислительных ограничений. Однако и здесь существуют барьеры: чем сложнее задача, тем больше ресурсов требуется для ее решения, что делает невозможным анализ всех альтернатив в разумные сроки. Например, в задачах, связанных с распознаванием образов или прогнозированием, система вынуждена использовать эвристики и приближенные методы, а не полный анализ всех возможных решений. Многие задачи, с которыми сталкивается искусственный интеллект, относятся к классу вычислительно сложных, что означает их экспоненциальный рост по количеству возможных решений или даже неразрешимость за полиномиальное время. Ограниченность вычислительных ресурсов в сочетании со сложностью реальных задач принятия решений делает достижение оптимальных решений практически невозможным. Одним из проявлений данной проблемы является комбинаторный взрыв: стремительное увеличение числа возможных вариантов, которое делает полный перебор решений неосуществимым даже для самых мощных вычислительных систем. В системах искусственного интеллекта это особенно заметно в задачах оптимизации и планирования, где экспоненциальное разрастание возможных комбинаций требует применения эвристических методов и приближенных алгоритмов. В результате ИИ приходится ограничивать область рассуждений некоторым *фреймом* (рамкой), что неизбежно упрощает картину мира, но предотвращает паралич от избыточных данных.

Концепция *ограниченной оптимальности* (*bounded optimality*), сформулированная Стюартом Расселом, предполагает, что оптимальным является не агент, действующий в абстрактных условиях, а тот, который достигает наилучшего возможного результата в пределах заданных ресурсных ограничений. Согласно его определению, «ограниченная оптимальность расширяет традиционное понимание рациональности, вводя в него эксплицит-

ный учёт ограниченных вычислительных ресурсов реальных агентов» [6, с. 13]. Данные представления легли в основу ряда исследований по рациональному мета-рассуждению, когда интеллектуальные агенты выделяют часть ресурсов не только на само решение задачи, но и на оценку эффективности дальнейшего поиска (например, когда целесообразно прекратить поиск решения и перейти к его реализации). Таким образом, от Саймона берет начало линия исследований, рассматривающая интеллектуального агента как рационально действующего в пределах своих возможностей, а не в абсолютном выражении. Сегодня практически все успешные системы ИИ в том или ином виде реализуют принципы ограниченной рациональности, срезая неперспективные варианты из-за ресурсных ограничений.

Однако ограниченность ИИ определяется не только техническими факторами, но и фундаментальными математическими принципами. Еще А. Тьюринг доказал существование неразрешимых задач, одной из которых является проблема останковки – задача, которую никакая универсальная вычислительная машина не способна решить [11; 12]. Этот факт имеет принципиальное значение: у любой формальной системы интеллекта существуют границы, обусловленные не только доступными ресурсами, но и самими законами вычислимости. Следовательно, ограничения искусственного интеллекта – не просто следствие недостаточных вычислительных мощностей, а фундаментальное свойство любой алгоритмической системы. Так же принципиальным ограничением являются фундаментальные барьеры, выявленные теоремами К. Гёделя о неполноте. Если в любой достаточно мощной формальной системе существуют утверждения, которые невозможно ни доказать, ни опровергнуть, то сходные проблемы могут возникать и в системах с искусственным интеллектом [10, с 265–266, с. 276]. Все это означает, что определенные классы задач останутся принципиально нерешаемыми для машинных алгоритмов, независимо от их сложности и вычислительных мощностей.

Можно сказать, что *ограниченная оптимальность* представляет собой компромисс между точностью и затратами: она не требует исчерпывающего анализа всех возможных решений, но гарантирует нахождение приемлемого результата в разумные сроки. Данный подход особенно актуален для реальных вычислительных систем,

которые должны функционировать в условиях ограниченного времени, памяти и вычислительных мощностей. Но даже этот уровень рациональности не всегда достижим, поскольку алгоритмы могут сталкиваться с ресурсными ограничениями, а вычислительная сложность некоторых задач делает поиск удовлетворительного решения невозможным. Современные интеллектуальные системы проектируются с учетом ограниченной рациональности: признается, что идеально рациональный алгоритм часто неосуществим, поэтому алгоритмы должны быть эффективными в пределах доступных ресурсов. Например, в системах планирования и поиска решений используются эвристики, отсечение нерелевантных вариантов и «*anytime-алгоритмы*», которые могут прерывать вычисление при достижении приемлемого решения.

Поскольку идеальная рациональность остается недостижимой, вычислительная рациональность ограничена ресурсами, а даже ограниченная оптимальность не всегда возможна, то наиболее реалистичной и практически применимой становится *ограниченная рациональность*, которая позволяет адаптировать искусственный интеллект к реальным условиям и существующим ограничениям. В этой модели агент принимает решения не на основе поиска глобально оптимального решения, а в рамках доступных возможностей и ограничений. Другими словами, система выбирает удовлетворяющее решение (*satisficing*) вместо поиска абсолютного оптимума. Такой подход широко применяется в алгоритмах искусственного интеллекта, особенно в многокритериальном принятии решений, где важно найти баланс между качеством решения и затратами на его вычисление.

Концепция *ограниченной рациональности* также играет ключевую роль в междисциплинарных исследованиях, особенно в изучении взаимодействия человека и искусственного интеллекта. Внедрение ИИ ставит перед исследователями ряд этических проблем, связанных с распределением ответственности за ошибки искусственного интеллекта. При таких обстоятельствах требуется определить, на ком лежит ответственность: на разработчиках, операторах или самой системе. Значимой задачей становится достижение оптимального соотношения между автономностью алгоритмов и необходимостью человеческого контроля, что предоставит возможность минимизировать вмешательство и в то

же время предотвратить возможные риски. Так, современные исследователи, в частности Л. Флориди, подчеркивают значимость концепции гибридного интеллекта, который предполагает сочетание машинных алгоритмов с человеческой интерпретацией и корректировкой [4, с. 163–165].

Одним из центральных аспектов данной проблематики является роль «человека в контуре» (*human-in-the-loop*), т.е. участие человека в принятии критически важных решений для минимизации рисков. Поскольку искусственный интеллект не обладает способностью к сознательному восприятию информации, то его следует рассматривать не как самостоятельного агента, а как инструмент, чья эффективность зависит от взаимодействия с человеком. Выделяются три основные модели такого взаимодействия. Первая модель обозначается как «человек над контуром» (*human-on-the-loop*) и предполагает, что человек контролирует решения ИИ и может вмешаться при необходимости. Вторая, «человек в контуре» (*human-in-the-loop*), предусматривает активное участие человека в процессе принятия решений, а также контроль и корректировку результатов работы алгоритмов. Третья модель, «человек вне контура» (*human-out-of-the-loop*), описывает ситуацию, когда система функционирует полностью автономно, без вмешательства человека, что повышает риски, особенно в критически значимых сферах.

По мере развития искусственного интеллекта важно не только совершенствовать алгоритмы, но и придавать большое значение философски взвешенному проектированию. Нужно помнить, что исключение человека из цикла принятия решений не снимает с него ответственности, а лишь скрывает её и затрудняет контроль. В ситуациях, где последствия носят выраженный моральный характер (к примеру, распределение медицинских ресурсов, определение виновности в судопроизводстве или поведение автономного автомобиля при угрозе ДТП), ответственность должна оставаться за человеком. Такой подход основывается на принципиальных отличиях между вычислительными процессами и человеческим опытом, включающим эмпатию, совесть и совокупное жизненное знание. В качестве иллюстрации можно привести следующие примеры. ИИ активно применяется в медицинской диагностике, однако его выводы требуют проверки врачом. Так, ал-

горитмы Google Health в 2021 году продемонстрировали 94% точности при выявлении рака молочной железы, но не распознавали редкие формы опухолей, которые специалисты определяли лучше [5]. Данный факт подтверждает необходимость гибридного подхода, при котором врач оценивает результаты ИИ перед постановкой диагноза. Похожие проблемы наблюдаются в сфере беспилотного транспорта. В 2018 году автомобиль Uber не распознал пешехода с велосипедом в темное время суток, что привело к смертельному ДТП [9]. В настоящее время автомобильные компании, включая Tesla, внедряют модель *human-on-the-loop*, требующую обязательного контроля со стороны водителя даже при активированном автопилоте. Современные исследования свидетельствуют о необходимости рассматривать искусственный интеллект не как полностью автономную систему, а как инструмент, работающий в тесном взаимодействии с человеком.

Благодаря коллективным усилиям исследователей и практиков, стремящихся повысить эффективность совместной работы человека и искусственного интеллекта, был разработан подход *Human-AI Teams*, целью которого является оптимизация взаимодействия человека и технологий, минимизация рисков автоматизированных решений и повышение качества результатов в различных сферах деятельности. Сочетание машинной аналитики и человеческого опыта открывает новые перспективы в медицине, транспорте, экономике и других критически важных сферах. В ближайшие годы ожидается совершенствование технологий адаптивного ИИ, способного не только выполнять поставленные задачи, но и постепенно обучаться взаимодействию с человеком, учитывая его предпочтения и стиль работы. Поэтому концепция ограниченной рациональности и модель «человека в контуре» становятся неотъемлемой частью развития искусственного интеллекта, обеспечивая баланс между эффективностью и этической ответственностью. При этом такое участие невозможно без обращения к концепции ситуационной рациональности, согласно которой принятие решений не может быть полностью алгоритмизировано, ведь каждая ситуация содержит в себе элементы неповторимости и открытости. В отличие от автономных машин, человек не просто исполняет функцию контроля или вмешательства, а привносит в процесс решения ценностный и интерпретативный компонент. Это означает,

что рациональность в модели *human-in-the-loop* не сводится к сугубо инструментальному мышлению: она опирается на способность субъекта учитывать контекст, неопределённость, моральные последствия и уникальность каждой ситуации. Человек в контуре не только оценивает работу алгоритма, но и обнаруживает контекстуальные нюансы, недоступные машине, работающей исключительно на основе статистических зависимостей. Его присутствие восстанавливает связь между рациональностью и практической мудростью (*phronesis*), то есть пониманием конкретной жизненной ситуации, в которую вписано каждое решение.

Таким образом, модель *human-in-the-loop* позволяет не только повысить безопасность и адаптивность систем, но и сохранить человеческое измерение решений в условиях цифровой рациональности. Участие человека в контуре представляет собой философски обоснованное требование, подчёркивающее важность интеграции рационального выбора с экзистенциальной, этической и социокультурной обусловленностью человеческих поступков. Это, в свою очередь, свидетельствует о необходимости согласования логики алгоритмических процедур с «логикой жизни» – сопряжения формальной рациональности и того, что философия называет «рациональностью второго порядка», предполагающей способность к саморефлексии и обоснованию не только результатов, но и исходных предпосылок выбора.

Заключение и выводы

В заключение еще раз подчеркнем, что рациональность искусственного интеллекта не может определяться идеальной моделью. Ограниченность вычислительных ресурсов и высокая сложность решаемых задач предопределяют необходимость обращения к концепции ограниченной рациональности. В этой связи применение концепции ограниченной рациональности становится необходимым условием для эффективного функционирования интеллектуальных систем. Их развитие требует сбалансированного сочетания формальной логики, доступных вычислительных мощностей и способности адаптироваться к динамично изменяющимся условиям окружающей среды. Перспективными направ-

лениями дальнейших исследований являются углубленный философский и методологический анализ гибридных моделей рациональности, объединяющих эвристические и формально-логические подходы. Важное значение также приобретает изучение этических последствий и социальных аспектов интеграции ИИ, что открывает широкие возможности для междисциплинарных исследований и практического применения полученных результатов. На основе анализа различных моделей рациональности установлено, что концепция «человека в контуре» (*human-in-the-loop*) играет ключевую роль в обеспечении контроля и корректировки решений, принимаемых искусственным интеллектом. Таким образом, ограниченная рациональность не только задает технические границы работы интеллектуальных систем, но и формирует концептуальный фундамент для разработки алгоритмов, учитывающих не только вычислительные, но и социальные и этические ограничения.

Литература

1. *Бабич М.Ю.* Понятие рационального агента и многоагентные системы // Проблемы информатики в образовании, управлении, экономике и технике: сб. статей XVII Междунар. науч.-техн. конф. Пенза: ПДЗ, 2017. С. 11–16.

2. *Елхова О.И.* Метрики феноменологического виртуального опыта // Вестник Российского университета дружбы народов. Серия: Философия. 2024. Т. 28. № 4. С. 997–1013.

3. *Елхова О.И., Кудряшев А.Ф.* Современные вызовы информационно-коммуникационных технологий // Вестник Самарского государственного технического университета. Серия: Философия. 2024. Т. 6. № 3. С. 27–34.

4. *Allo P.* (ed.). *Putting Information First: Luciano Floridi and the Philosophy of Information.* Wiley-Blackwell, 2011. 208 p.

5. *Forster V.* Artificial Intelligence Improves Breast Cancer Diagnosis [Электронный ресурс] / V. Forster. Режим доступа: <https://www.forbes.com/sites/victoriaforster/2025/01/13/artificial-intelligence-improves-breast-cancer-diagnosis/> (дата обращения: 15.03.2025).

6. *Russell S.* Rationality and Intelligence: A Brief Update // In: Müller V.C. (ed.). *Fundamental Issues of Artificial Intelligence.* Cham: Springer, 2016. P. 7–28.

7. *Russell S., Norvig P.* Artificial Intelligence: A Modern Approach. Pearson Series in Artificial Intelligence, 2022. 1115 p.

8. *Simon H.* Models of Man: Social and Rational. New York: John Wiley and Sons, Inc., 1957. 287 p. URL: <https://archive.org/details/modelsofman0000herb> (дата обращения: 15.01.2025).

9. The Guardian. Uber self-driving car kills woman in Arizona [Электронный ресурс] // The Guardian. Режим доступа: <https://www.theguardian.com/technology/2018/mar/19/uber-self-driving-car-kills-woman-arizona-tempe> (дата обращения: 15.03.2025).

10. *Tourlakis G.* Gödel's First Incompleteness Theorem via the Halting Problem // *Computability*. Springer, 2022. P. 265–280.

11. *Turing A.* Computing Machinery and Intelligence // *Mind*. 1950. Vol. 49. P. 433–460.

12. *Turing A.* On Computable Numbers, with an Application to the Entscheidungsproblem // *Proceedings of the London Mathematical Society*. 1936. Vol. 42. P. 230–265.

References

1. *Babich M.Yu.* Ponyatie ratsional'nogo agenta i mnogoagentnye sistemy [The Concept of a Rational Agent and Multi-Agent Systems] // *Problemy informatiki v obrazovanii, upravlenii, ekonomike i tekhnike: sb. statei XVII Mezhdunar. nauch.-tekhn. konf.* [Problems of Informatics in Education, Management, Economy, and Engineering: Proceedings of the 17th International Scientific and Technical Conference]. Penza: PDZ, 2017. P. 11–16. (In Russ.)

2. *Elkova O.I.* Metriki fenomenologicheskogo virtual'nogo opyta [Metrics of Phenomenological Virtual Experience] // *Vestnik Rossiiskogo universiteta druzhby narodov. Seriya: Filosofiya.* [RUDN Journal of Philosophy]. 2024. Vol. 28. No. 4. P. 997–1013. URL: <https://doi.org/10.22363/2313-2302-2024-28-4-997-1013>. (In Russ.)

3. *Elkova O.I., Kudryashev A.F.* Sovremennye vyzovy informatsionno-kommunikatsionnykh tekhnologii [Modern Challenges of Information and Communication Technologies] // *Vestnik Samarskogo gosudarstvennogo tekhnicheskogo universiteta. Seriya: Filosofiya.* [Bulletin of Samara State Technical University. Series: Philosophy]. 2024. Vol. 6. No. 3. P. 27–34. (In Russ.)

4. *Allo P.* (ed.). *Putting Information First: Luciano Floridi and the Philosophy of Information*. Wiley-Blackwell, 2011. 208 p.

5. *Forster V.* Artificial Intelligence Improves Breast Cancer Diagnosis [Electronic resource] / V. Forster. Available at: <https://www.forbes.com/sites/victoriaforster/2025/01/13/artificial-intelligence-improves-breast-cancer-diagnosis/> (accessed: 15.03.2025).

6. *Russell S.* Rationality and Intelligence: A Brief Update // In: Müller V.C. (ed.). *Fundamental Issues of Artificial Intelligence*. Cham: Springer, 2016. P. 7–28.

7. *Russell S., Norvig P.* *Artificial Intelligence: A Modern Approach*. Pearson Series in Artificial Intelligence, 2022. 1115 p.

8. *Simon H.* *Models of Man: Social and Rational*. New York: John Wiley and Sons, Inc., 1957. 287 p. URL: <https://archive.org/details/modelsofman0000herb> (accessed: 15.01.2025).

9. The Guardian. Uber self-driving car kills woman in Arizona [Electronic resource] // The Guardian. Available at: <https://www.theguardian.com/technology/2018/mar/19/uber-self-driving-car-kills-woman-arizona-tempe> (accessed: 15.03.2025).

10. *Tourlakis G.* Gödel's First Incompleteness Theorem via the Halting Problem // *Computability*. Springer, 2022. P. 265–280.

11. *Turing A.* Computing Machinery and Intelligence // *Mind*. 1950. Vol. 49. P. 433–460.

12. *Turing A.* On Computable Numbers, with an Application to the Entscheidungsproblem // *Proceedings of the London Mathematical Society*. 1936. Vol. 42. P. 230–265.

Информация об авторе

Елхова Оксана Игоревна – Уфимский университет науки и технологий, доктор философских наук, доцент, профессор кафедры философии и культурологии (Республика Башкортостан, 450076, Уфа, ул. Заки Валиди, 32).

oxana-elkhova@yandex.ru

Information about the author

Elkhova Oxana Igorevna – Ufa University of Science and Technology, Doctor of Philosophy, docent, Professor at the Department of Philosophy and Cultural Studies (Republic of Bashkortostan, 450076, Ufa, Zaki Validi Street, 32).

oxana-elkhova@yandex.ru

Дата поступления 10.12.2024

Принята к печати 11.12.2025