

ФИЛОСОФИЯ

Научная статья
УДК 165
doi: 10.17223/15617793/512/7

Перспективы исследования метапознания для развития естественного и искусственного интеллектов

Юлия Владимировна Логиновская¹

¹ Национальный исследовательский Томский государственный университет, Томск, Россия, urucel@yandex.ru

Аннотация. Представлены результаты изучения философско-методологических аспектов метапознания. Раскрыта взаимосвязь метапознания с эпистемологией в целом и с личностной эпистемологией. Проанализирована триада метапознания (метазнание, метакогнитивные переживания, метакогнитивная регуляция) в контексте развития естественного и искусственного интеллекта. Показаны возможности и ограничения метапознания для искусственного интеллекта. Сделан вывод о перспективах дальнейшего исследования метапознания.

Ключевые слова: метапознание, метакогнитивная стратегия, личностная эпистемология, рефлексивное знание, неявное знание, мудрость

Источник финансирования: исследование выполнено в рамках Программы стратегического академического лидерства «Приоритет 2030».

Для цитирования: Логиновская Ю.В. Перспективы исследования метапознания для развития естественного и искусственного интеллектов // Вестник Томского государственного университета. 2025. № 512. С. 61–69. doi: 10.17223/15617793/512/7

Original article
doi: 10.17223/15617793/512/7

Prospects of metacognition research for the development of natural and artificial intelligence

Yuliya V. Loginovskaya¹

¹ National Research Tomsk State University, Tomsk, Russian Federation, urucel@yandex.ru

Abstract. Today the research on metacognitive processes is mostly conducted in practice-oriented scientific fields: psychology, pedagogy, and the development of artificial intelligence systems. But philosophy (in particular, epistemology) has the theoretical potential to explicate metacognition. Therefore, the aim of the study is to identify the philosophical and methodological foundations of metacognition, as well as promising areas of research and development of metacognitive strategies for the development of natural intelligence and artificial intelligence. Based on a philosophical and methodological analysis, the article provides a philosophical interpretation of the concept of metacognition and its elements: meta-knowledge, metacognitive experience, and metacognitive regulation. The article reveals the ontological and epistemological prerequisites for studying metacognitive processes and draws analogies with relevant topics in philosophy. The author clarifies the relationship between knowledge and information, cautions against reducing knowledge to data or information, and substantiates the relationship between meta-knowledge and wisdom as characteristics of the maturity of consciousness. The research revealed three aspects of the study of metacognition: reflexive, epistemological, and ontological. As a result, the author highlights the role of reflection in the formation of metacognition in natural intelligence, taking into account the peculiarities of personal epistemology development. Then the author asserts the formation of a metacognitive subject as an observer with metareflexion. At the same time, the role of implicit knowledge is substantiated on the basis of the 4E-approach in cognition (embodied, embedded, enacted, extended). This leads to the understanding of the integrity of the cognitive situation, in which both bodily experience and subjective experiences are important, as well as skills of self-organization and self-regulation being the part of the metacognitive triad. The study of the use of metacognitive strategies in the development of artificial intelligence revealed the possibilities and limitations of metacognition for artificial systems related to their "metacognitive myopia". The author makes the conclusion about a tendency of interdisciplinary research combining philosophy and theories of consciousness, personal epistemology and metacognition to transform the process of the "natural – artificial intelligence" system development, in which human intelligence plays a leading role, and artificial intelligence plays an auxiliary role. The article ends with the conclusion about the inevitable development of metacognition as an individual metacognitive competence and the relevance of the formation of a metacognitive strategic type of thinking.

Keywords: metacognition, metacognitive experience, metacognitive strategy, personal epistemology, reflexive knowledge, implicit knowledge, wisdom

Financial support: The study was carried out within the framework of the Strategic Academic Leadership Programme "Priority 2030".

For citation: Loginovskaya, Yu.V. (2025) Prospects of metacognition research for the development of natural and artificial intelligence. *Vestnik Tomskogo gosudarstvennogo universiteta – Tomsk State University Journal*. 512. pp. 61–69. (In Russian). doi: 10.17223/15617793/512/7

В ситуации, когда научные дисциплины и технологии сталкиваются с новыми вызовами, не решаемыми старыми способами, на помощь приходит философия. Ускоряющиеся процессы разработки и применения искусственного интеллекта заставляют исследователей не только заново разграничивать и пересматривать часто используемые понятия, такие как «информация», «знание», «познание», но и переосмысливать процессы производства и передачи знания. Так, с помощью своеобразной рефлексивной надстройки с приставкой «мета-», в научно-исследовательской литературе и современных исследовательских проектах обозначается метауровень организации познания с помощью описания метакогнитивных процессов (метапознания, метакогнитивного опыта, метакогнитивных стратегий, метапонимания, метапамяти и т.д.). Насколько оправданным и перспективным является развертывание нового направления, можно понять, лишь разобравшись в его содержании. Поэтому целью настоящего исследования стало выявление философско-методологических оснований метапознания и перспективных направлений исследования и разработки метакогнитивных стратегий для носителей естественного интеллекта и искусственного интеллекта (ИИ).

Термин «метакогнитивные процессы» является зонтичным и включает в себя целый комплекс рефлексивных, организационных, регулятивных и даже телеологических аспектов метапознания [1]. Метапознание сегодня является в первую очередь предметом исследования метакогнитивизма – междисциплинарного направления, включающего помимо спектра когнитивных наук такие отрасли, как философия сознания, инженерия знаний, педагогическая психология и другие. Однако если в философии, в частности, в теории познания и эпистемологии, большее внимание сегодня уделяется собственно понятию и структуре знания, а также подходам к созданию нового знания [2], то метакогнитивные процессы, связанные с наблюдением, рефлексией, оценкой и организацией познания, чаще соотносятся с более узконаправленными и практико-ориентированными дисциплинами: педагогическая психология, когнитивные науки, философия образования и др.

Согласно автору термина «метапознание» Дж. Флейвеллу (1977), данный феномен тесно связан с самосознанием, а именно со способностью человека регулировать свою когнитивную деятельность [3]. Метакогнитивная компетенция опирается на «знание о собственных когнитивных процессах, их продуктах и обо всем, что связано со свойствами, имеющими отношение к изучению информации и данных» или, если выразиться кратко, является «знанием второго порядка». Флейвелл рассматривал явление метапознания

в двух измерениях: в первом рассматривается *метазнание* как *декларативный компонент* метапознания (личные знания или убеждения о собственных когнитивных процессах и факторах, способствующих или препятствующих получению знания, а также осознание ресурсов для познания). Другими словами, декларативное метазнание образует «*эпистемологическую осведомленность*» в купе с навыками самопонимания, самоэмпатии и саморепрезентацией, т.е. предполагает достаточно высокий уровень сознательности, характеризующийся наличием «внутреннего наблюдателя». Второй – *процедурный* – компонент метапознания затрагивает активный контроль и самостоятельное регулирование процесса познания для достижения определенного познавательного результата.

В современном представлении в структуре метапознания выделяются три основных составляющих: метакогнитивное знание, метакогнитивное переживание и метакогнитивная регуляция. Метакогнитивное знание можно определить как рефлексивное знание о когнитивных и личностных детерминантах познания, которое объединяет в себе следующие аспекты: *декларативное* знание о своих ресурсах и возможностях; *процедурное* знание цели познания и процессов, необходимых для решения проблем; *знание условий*, при которых можно генерировать, передавать и применять знание; *знание об особенностях познания на своем этапе развития* (developmental knowledge) [4].

В философско-методологическом ключе метазнание охватывает онтологические и гносеологические установки, понимание закономерностей познавательного процесса, осведомленность о своем незнании и о методах познания; именно поэтому можно утверждать, что способность к метапознанию нарабатывается при изучении философии и критического мышления. Нарботанное метапознание выводит ученых, исследователей к формированию собственной метапознавательной позиции в виде авторской разработанной методологии, которая может стать метазнанием для других [5. С. 17–18].

Метакогнитивные переживания или чувства весьма многообразны и связаны с ощущениями легкости или трудности обучения, опытом самоэффективности, уверенностью в способности решить задачу или состоянием неопределенности при встрече с познавательной задачей, что обозначается такими состояниями, как «чувство знания», «чувство близости к решению» и «чувство неуспешности в познании» и т.п. [6]. Схожее описание метакогнитивных чувств встречается у Б. Рассела: «Вызывает ли предложение чувство знания или чувство сомнения, зависит от того, оставляет ли оно открытыми альтернативы, призывающие к различным действиям и эмоциям».

Каждая дизъюнкция, которая не является логически исчерпывающей (то есть не имеет вид «А» или не-«А»), даёт некоторую информацию, если она истинна, но информация может оставить нас настолько сомневающимися в том, что делать, что она ощущается как незнание» [7]. В максимальном обобщении метакогнитивные чувства можно свести к интуиции как элементу неявного знания. Метакогнитивные переживания обуславливают субъективные особенности метапознания, которые напрямую связаны с неявным (или контекстуальным) знанием.

Наличие и роль такого интуитивного, контекстуального знания и опыта в метапознании можно назвать основным «камнем преткновения» для развития искусственного интеллекта, не обладающего интуицией как опытом глубинных, неявных связей знания с реальностью человеческого мира. Метакогнитивное переживание по сути является *квалиа* или *сознательным опытом*, в терминах Д. Чалмерса [8]. Соответственно, перспективно рассматривать данный компонент метапознания как атрибут телесно укорененного знания с соответствующей рефлексивной практикой (осознанием) и «проживанием» знания, встраиванием его в структуру жизненных связей. Мы исходим из того, что метапознание в целом определяется субъектностью, т.е. зрелостью сознания и личности, которая характеризуется наличием внутреннего импульса к действию и созидательной, организующей активность сознания. Конечной целью метапознания является выработка собственной эффективной метакогнитивной стратегии для успешного овладения знанием, а для этого должен существовать «сознательный метастратегический уровень мышления» [9].

Поскольку уровень доступности информации и объёма данных на сегодняшний день более чем высокий, особую значимость в метапознании приобретает также его третий компонент – саморегуляция как способность управлять процессом собственного когнитивного развития на основе знания о знании [6]. Саморегуляция в контексте метапознания тесно связана, во-первых, с умением взаимодействовать с информационным потоком, сознательно ограничивая поиск информации и выбирая релевантные источники познания; во-вторых, с самокоррекцией – выявлением и исправлением пробелов в своем познании, отображением качества своего познания на основе некой нормы. Внедрение метакогнитивной практики как саморегуляции при обучении. В исследовательской литературе отмечается, что по мере того, как вырастает значение метакогнитивных стратегий и укрепляется связь между самосознанием и когнитивной реакцией, метапознание в большей мере с саморегуляцией, нацеленной в первую очередь на когнитивную сферу, а не на психосоциальную или поведенческую [10]. Однако среди представителей когнитивных наук наблюдаются расхождения в понимании компонентов метапознания и его соотношения с саморегуляцией.

С философской точки зрения саморегуляция рассматривается как подчиненный элемент стратегического процесса метапознания, а регулятивное начало отводится мышлению как источнику метакогнитивных

компетенций, обеспечивающих возможность сознательной организации собственного познавательного процесса. «Эпистемический мониторинг» (саморефлексия и самопроверка результата познания) поддерживает сознательную регуляцию познавательного процесса, однако для его осуществления требуется способность внутреннего отстранения от познавательной ситуации и взгляда на неё с позиции «наблюдателя», что подразумевает определенную степень развития сознания. Возможно, именно на такой процесс указывает М. Фуко: «В современном когито речь идет о том, чтобы понять все значение того промежутка, который одновременно отделяет и вновь соединяет сознающую себя мысль и ту ее часть, которая укореняется в нем самом» [11. С. 146].

Метакогнитивная регуляция, третья составляющая в структуре метапознания, содержит такие элементы, как мониторинг, критическая оценка, контроль и планирование когнитивных процессов, а также включает стратегии исправления ошибок понимания и управление информацией посредством знания. В данном процессе учитываются не только внутренние механизмы самоорганизации и способы мышления, но и непосредственное влияние социокультурной среды и взаимодействий как стимулирующего фактора метапознания. С точки зрения последователей саморегулируемого обучения, саморегуляция понимается как сверхупорядоченное проявление метапознания [10]. В саморегулируемом обучении на первый план выходит организационная составляющая метакогнитивности как способность выбирать, координировать и комбинировать когнитивные стратегии, связанные с переводом информации в форму знания, его концептуализации и интериоризации [12].

Согласно Вайнштейну и Майеру, три типа когнитивных стратегий (повторение, разработка и организация) связаны с последовательными процессами кодирования новой информации и включения их в рабочую память, сохранением информации в долговременной памяти, структурированием и синтезом информации [13]. Но, кроме стратегий обработки информации, существуют и другие типы познавательных стратегий, затрагивающие иные аспекты актуализации знания (то же контекстуальное или неявное знание) с целью более глубокого его усвоения и использования. Помимо этого, для саморегуляции важно когнитивное моделирование – способность уметь моделировать метапознание, формируя способность к самопониманию, самодисциплине и самоорганизации (ключевые *self-навыки*) вкпе с наблюдением за метакомпетенциями других.

Особенность метапознания в том, что оно может быть как индивидуальным, так и коллективным. По мере совершенствования метакогнитивных навыков также повышаются шансы на успешное решение командных задач и проектов: «Исследования показывают, что метапознание в социальных условиях может помочь смягчить некоторые проблемы индивидуального метапознания и способствовать ему, поскольку каждый член команды сталкивается с метакогнитивной саморефлексией других, получает обратную связь

и рассматривает стратегии саморегуляции других» [14. Р. 151]. Таким образом, регулятивная триада «метапознание, саморегуляция и саморегулируемое обучение» становится методологическим ориентиром для более эффективной организации образовательного процесса и формирования метакогнитивной компетентности самого обучающегося. Но при этом под вопросом остается соотношение и значение каждого из аспектов триады для производства знания. Комплексная разработка каждого из выделенных аспектов метапознания позволяет более эффективно организовать, структурировать и корректировать когнитивную ситуацию и включить в нее творческий аспект.

Информация и знание в метапознании. В развитии концепции метапознания как для естественного, так и для искусственного интеллекта важным является вопрос разграничения информации и знания. Многие определения знания базируются на понятии информации и используются в словарях как тождественные или взаимозаменяемые, а базы данных в компьютерных науках часто называются базами знаний, хотя они взаимодействуют только с информацией. Например, взаимозависимость понятий «информация», «данные», «знание» разделяются Р. Буркхардом, согласно которому «данные – это факты, информация – это интерпретация фактов, знание – персонализированная информация» [15].

По справедливому замечанию С.Н. Костроминой, Д.С. Гнедых, подмена понятия «информация» понятием «знание» зачастую встречается в образовательной среде и в процессе «обмена знаниями» в профессиональных группах, где в процессе коммуникации, по сути, происходит «интеграция новой информации в смысловую систему субъекта, процедуры понимания, усвоения, применения и т.д.» [16]. В контексте современных когнитивных вызовов ценность метапознания будет заключаться в наработке способности осознанного выбора и различения информации как основания, на котором строится знание, однако личностное знание не сводится только к вербализованной части, а имеет более широкий спектр проявлений и также влияет на успешность познавательной ситуации.

Редукция понятия знания к процессу обработки и усвоения информации, которое происходит в процессе разработки систем искусственного интеллекта, можно назвать непродуктивным, поскольку: во-первых, не учитывается организационный и созидательный потенциал знания, а во-вторых, игнорируется аспект владения знанием, уровень компетентности, который больше относится к категории умения, чем осведомленности. Частично такую редукцию снимает разграничение декларативного «знания, что» и процедурного «знания, как». Согласно А.Л. Никифорову, знание «воплощается (конденсируется) в смысле наших слов» [17], т.е. знание можно рассматривать как динамическую смысловую интегративную единицу, оформленную вербализованно или образно и адекватно описывающую реальность. Активная метапознавательная рефлексия и задает эту динамику.

В философском понимании знание можно определять не как некую единицу, а единство форм человеческого опыта в целостном бытии. Так, единство знания,

согласно К.С. Пригову, отражает состояние реального мира: «Принципиальный факт множественности единого мира на всех уровнях (реальности, знаков и знаний) предполагает возможность знания только как единства знания» [18. С. 128]. Качественное отличие знания от информации состоит в единстве знания, его включенности в личностный опыт и в структуру социальных взаимосвязей: «В пределе мы действительно знаем (т.е. понимаем смысл информации), если сумели включить данную информацию в систему всех знаний, наличествующих в человечестве» [16].

Переход от информации к знанию на практике – это конструктивный процесс формирования знания, заключающийся в последовательной интеграции новой информации в смысловую систему субъекта и включающий процедуры понимания, усвоения, применения [17], а также и в социальную структуру взаимосвязей. Во всех познавательных процедурах надстройка в виде метапознания обеспечивает эффективность процесса включения знания в структуру сознания субъекта.

Перспективным видится исследование понятия мудрости в контексте метапознания. В известной иерархической модели Р. Акоффа «данные-информация-знание-мудрость» (DIKW) характеристики знания, описанные выше, будут соотноситься с верхним уровнем пирамиды, а именно проявлением мудрости как «знания правильных вещей», умения использовать знание во благо и возможности принимать взвешенные, здравые решения на основе знания [19]. Мудрость, понимаемая как форма общественного сознания или как целостная форма познания человека и мира, содержит регулятивный потенциал [20], который, на наш взгляд, заключается в целесообразности выбора стратегий получения и применения знания с учетом этических и жизнеспасающих критериев. С философской точки зрения связь метапознания и мудрости может быть обозначена в контексте понятия «логоса», т.е. обретения внутреннего нормирующего, упорядочивающего фактора саморазвития. Мудрость предполагает определенную зрелость сознания и напрямую относится к метапознанию. Можно утверждать, что метапознание без постижения мудрости вряд ли достижимо.

Очевидно, что основной проблемой использования метакогнитивных стратегий является недоразвитость как неспособность просчитать долгосрочные перспективы или трезво оценить необходимые шаги для получения результата со стороны познающего. Соответственно, обучение метапознанию обусловлено и сопровождается внутренним «взрослением» человека, повышением ответственности за познание и его результат; а в своем высшем проявлении – наработкой мудрости как высшей формы сознания и формированием экспертной системы знаний. Рассматриваемое в таком ракурсе метапознание – это способность осознавать свою недоразвитость в какой-либо области знания, а также понимание необходимости непрерывной познавательной активности и принятие ответственности за качество производимого и передаваемого знания.

Метапознание и эпистемология. Большинство практических вопросов в области метапознания решаются психологическими и когнитивистскими подходами.

Сегодня в педагогической теории постулируется успешность процесса познания в ее прямой связи с метапознанием и, соответственно, разрабатываются стратегии обучения. Философский же взгляд не ограничивается операциональными разработками метакогнитивизма, а позволяет содержательно дополнить понимание метапознания новым смысловым измерением. Чем же здесь может быть полезной философия и такой ее раздел, как эпистемология? Во-первых, раскрыть суть «знания о знании» или метапознания, во-вторых, оптимизировать работу со знанием как с динамическим явлением через определение и разработку критериев мета навыков и метакомпетенций.

Согласно В.А. Лекторскому, задача эпистемологии двуединая: сформулировать стандарты различения знания, незнания и псевдознания (прескриптивная) и вместе с тем описать реальный успешный познавательный процесс (дескриптивная и объяснительная) [2. С. 8]. Рассмотрим три философско-методологических основания для исследования метапознания, связанных с философским рассмотрением знания, познания и сознания: рефлексивное, эпистемологическое и онтологическое.

Рассматриваемое через призму рефлексии, метапознание характеризуется рефлексивной абстракцией новых или существующих когнитивных структур, обращаясь к самому процессу познания и мышления [21]. Рефлексирующее сознание как базовая метасистема «метакогнитивной сферы личности» (согласно А.А. и А.В. Карповым) становится ключевым, интегративным параметром изучения метапознания, позволяющим фиксировать «рефлексивное расслоение» и проявляющим наиболее глубинные, сущностные закономерности и механизмы метакогнитивного плана [22. С. 91]. Именно осознание или рефлексия является пусковым механизмом формирования суждений или оценок, являющихся элементами знания. И наоборот, метапознание можно рассматривать как катализатор процессов эмоциональной регуляции и общей рефлексии, т.е. метапознание и саморегуляция являются взаимовыгодными, синергетическими механизмами. Перспективной можно назвать такую точку зрения, в которой метапознание не отождествляется с рефлексией, а концептуализируется как фундаментальная конструкция (метарефлексия), дополняющая рефлексивную активность сознания.

Перспективные предпосылки к эпистемологической разработке вопросов метапознания встречаются у отечественных исследователей. Так, М.А. Розов в своей эпистемологической концепции социальных эстафет в рамках деятельностного подхода говорит о содержании знания как формы социальной памяти и ведет речь о «надрефлексивной позиции», когда «исследователь рефлектирующей системы пытается как бы вывести себя за пределы этой системы и посмотреть на нее со стороны» [23]. Любая деятельность, в том числе познавательная и производственная, тесно связана с целеполагающей рефлексией и проектированием. Рефлексивные преобразования в ходе познания и творчества выражаются сменой целевых установок и проявляются разными актами осознания одного содержания

знания на теоретическом и практическом уровне. Метапознание, таким образом, выступает условием трансформации и развития знания благодаря рефлексивной «проекции» сознания.

Эпистемология больше внимания уделяет собственно определению, организации, объективным структурам и критериям научного знания, поэтому фокус её внимания лишь косвенно относится к метапознанию. Так, Э. Агацци выделяет в эпистемологии два аспекта: дескриптивный и нормативный. Нормативный аспект является «предварительным», дающим некое достаточное определение знания и уточняющим требования к знанию, а дескриптивный состоит в выяснении способов, условий и критериев получения знания [24]. По замечанию Э. Агацци, эпистемология развивается по принципу формирования петель обратной связи, причем её парадигмальная динамика оказывается больше связана с развитием дескриптивного аспекта, что не даёт выхода на метауровень, связанный с исследованием метакогнитивных процессов. Тем не менее перспективы включения эпистемологии в исследовательское поле метапознания кажутся оправданными. Нарботки эпистемологии могли бы существенно дополнить изучение метапознания, к примеру, методами трансцендентальной феноменологии как философии, сопутствующей «информационно-коммуникационной, а потом и когнитивной революции» для прояснения метакогнитивного опыта познания [25]. Но при этом стоит признать, что возможности трансцендентальной феноменологии в данном случае ограничены, поскольку отказ феноменологии от «внешней» стороны человеческой деятельности в пользу трансцендентальной интенциональности не позволяет объяснить другие аспекты метапознания, связанные с регуляцией познавательного аспекта.

Онтологически обоснованную идею конструктивного принципа понимания в формировании «знаниевых образований» высказывал М.К. Мамардашвили в своей естественно-исторической гносеологии, опирающейся на онтологическую укорененность знания в сознании. Нарбатывая способность метапознания, мы формируем, как сформулировал бы это М.К. Мамардашвили, «понимательную связь субъекта познания с уходящим в бесконечность миром знания»; и это «путь метасферы» сознания, в котором рождается и переживается опыт познавательного взаимодействия [26. С. 53]. В конкретном содержании знания помимо структурных и одновременных связей существуют, преобразуются и непрерывно обновляются генетические мыслительные связи, коренящиеся в предметной деятельности и обладающие внутренней историей истины. Для искусственного интеллекта именно этот момент знания остаётся недоступным в силу отсутствия субъектности и историчности, выражаясь словами М.К. Мамардашвили, той «временной глубины, складок и полости конечной протяженности вечной истины», образующей «разнопространственные мешочки», «многие глубины» нашей памяти [26. С. 52–53].

Очевидно, что гораздо проще, особенно для разработчиков искусственного интеллекта, свести знание к информационной составляющей (как было указано

выше) и не учитывать онтологическое измерение знания, поскольку оно требует особых методологических принципов усвоения и генерации знания, но, тем не менее без понимания онтологической укорененности знания невозможно будет выстроить и адекватное человеческой реальности метапознание. Рассудочные менталистские представления о знании ограничиваются лишь информационной структурой, в то время как реальные явления и продукты связаны с генеративными свойствами жизни и познания, что дает глубину познания: «...есть такие вещи, которые должны быть пережиты и заново установлены, непрестанно рождаясь в лоне бесконечных потенций. Важно происходящее внутри индивидуального переживания, во внутрь протянутом измерении эмпирического факта испытания» [26. С. 16].

Если принять во внимание тезис М. Мамардашвили о временно-пространственно протяженной форме существования знания, сознательное «размещение» знания в культурно-исторический, социальный, практический, ценностный и другие контексты, синтез знания в процессе познания будет наиболее эффективным для конструирования знания как реального продукта, применимого в общей системе «человек-мир».

В отношении сознательного наблюдения (метакогнитивного мониторинга) как необходимого условия знания возникает вопрос: возможно ли, что метапознание в конечном итоге приведёт к формированию некоего наблюдательного и контролирующего «органа» или «центра познания», контролирующего процесс познания? В качестве ответа можно привлечь понятие метапознающего субъекта как наблюдателя, обладающего метарефлексией, в концепции, выросшей из системной теории. Предпосылки к метапозиции в познании видим у Б. Рассела: «Наблюдатель, когда ему кажется, что он наблюдает за камнем, на самом деле, если верить физике, наблюдает за воздействием камня на самого себя» [7. Р. 15]. У Н. Лумана не встречается термин «метапознание», но, по сути, мониторинг знания представляет собой «наблюдение второго порядка», т.е. наблюдение наблюдателя на предмет того, что он может и что он не может видеть» [27. С. 161].

Развивая концепцию «наблюдателя второго порядка», отечественные философы В.И. Аршинов, В.Г. Буданов и Я.И. Свирский делают важный вывод о том, что наблюдение является операцией, воспроизводящей наблюдателя: «Наблюдатель демонстрирует все особенности системы, поскольку для наблюдения она должна быть способна производить и воспроизводить себя в окружающей среде. Любому наблюдению присущи познание и воля, то есть при наблюдении наблюдатель... упорядочивает и обменивается указаниями, которые он использует, чтобы вызвать свой мир, указывая и тем самым различая его» [28]. Метапознающий субъект является конструктором-проектировщиком и одновременно корректором своего познания, который способен мыслить личностно, голографически, дифференциально и при этом «мыслить в контексте рекурсивно-целостного процесса саморазличий, где внутреннее и внешнее топологически связаны» [28. С. 59].

Следовательно, одной из основных компетенций, связанных с метапознанием, можно назвать дифференциацию теоретического знания на основе различения контекстов (культурно-исторических, родовых, социальных, пространственно-временных и пр.), а в плане практического знания – способность к интроспективному осознанию реалий текущей ситуации, возможных вариантов выбора и требуемых решений базируется на умениях «само» – самосознания, самооценки и, соответственно, различения границ «своего» и «чужого», которые тесно связаны и с телесностью, вписанной в рекурсивный процесс взаимодействия с окружающей средой.

Метапознание и личная эпистемология. В англоязычной литературе метакогнитивные процессы рассматриваются в рамках направления, обозначаемого как «личная эпистемология», которое в большей мере можно отнести к педагогической психологии, нежели к философии [29, 30]. Согласно представителям этого подхода, основанного на идеях Л.С. Выготского, субъект познания вырабатывает «личную эпистемологию» на основе понимания самого процесса познания, конструирования и оценки знания с учетом личностных характеристик. При этом эпистемологическое (метакогнитивное) развитие начинается в позднем подростковом или юношеском возрасте с постепенно «развивающейся способностью координировать субъективную и объективную природу знания» [30. Р. 225].

Практической задачей личной эпистемологии становится формирование индивидуальных концепций знания и познания (эпистемических убеждений и теорий). Проблема источника знания в личной эпистемологии тесно связана с представлением о соотношении информации и знания: знание, содержащееся на каком-либо внешнем для человека носителе, по сути, является для него лишь «источником информации» или «предварительным знанием», на основе которого конструируется смысл и знание как «свое».

В рамках личностно-эпистемологического направления исследуются различные виды знания: знание (knowledge) как некое устойчивое образование и знание (knowing) как процесс – гибкое, динамичное явление; простое (изолированное) знание и сложное (взаимосвязанное с другими знаниями), как достоверное и явное образование и неявное (близкое к догадке) [29]. Процессуальное понимание знания и его формирования предполагает не только активное включение познавательных процессов, мыслительных операций (аналитико-синтетических, трансформационных и творческих), но и непрерывный контакт с миром для сверки актуальности и применимости знания, что составляет суть метакогнитивной регуляции.

Базовые различия видов и источников знания определяют характер метапознания, в котором большую роль играет неявное знание. С.А. Филипенко, раскрывая природу неявного знания естественного интеллекта в рамках подхода 4E-Cognition как (embodied – телесно-воплощенное, embedded – вписанное в среду, enacted – связанное с действием и extended – расширенное), показывает, что при познании конкретное действие индивида порождает особую когнитивную ситуацию, в которой познавательный акт, приводящий к

появлению нового знания, зависит и от телесного опыта (психофизических особенностей), и от предпосылок когнитивного акта в субъективной реальности переживаний, и от влияния внешней среды. В познавательном акте человек участвует целостно, всем своим жизненным опытом. Применяя подход 4E-Cognition к искусственному интеллекту, С.В. Филипенко указывает, во-первых, на то, что «в познавательной и творческой деятельности участвует не сам по себе искусственный интеллект, а целостная система «естественный интеллект – искусственный интеллект», а во-вторых, на ограничения искусственного интеллекта в плане многообразного, неявного телесного и социального опыта, сформированного у человека эволюционным путем [31. С. 108]. Согласимся, что в деятельности искусственного интеллекта корректнее говорить об обработке информации, а не продуцировании знания, поскольку информация становится знанием в результате наделения ее человеческим смыслом, включающим в себя вышеуказанные аспекты.

Таким образом, обозначенная тенденция исследовательского объединения философии и теорий сознания, личностной эпистемологии и метапознания для трансформации процесса познания достаточно перспективна в большей степени для развития системы «естественный – искусственный интеллект», в которой человеческий интеллект играет ведущую роль, а искусственный – вспомогательную.

Метапознание у систем ИИ. Но всё же возможен ли самостоятельно действующий метакогнитивный искусственный интеллект? В 2024 г. разработчиками систем искусственного интеллекта была анонсирована возможность обретения метапознания искусственным интеллектом, хотя принципы саморегулируемого обучения, основанного на механизме обратной связи, уже давно используются. Отсутствие самости, телесности и живой рекурсии, не ограниченной информационными потоками, у искусственных интеллектуальных систем, соответственно, вносит ограничения в области метапознания. Проблема метапознания у искусственного интеллекта перекликается с вопросами его концептуальных возможностей (понимания, символического познания и т.п.). Среди исследователей взгляды различаются: встречается как скептическое отношение к способности понимания у генеративных моделей [32], так и перспективное видение обучения моделей на все более сложных эвристиках и паттернах человеческого интеллекта [33]. Фокус исследователей при обучении систем искусственного интеллекта сегодня состоит не в запоминании, а в моделировании, отображении взаимосвязей и закономерностей в данных, а это требует уже метапознавательной активности. Влияние окружающей среды в виде запросов и команд пользователей (для искусственного интеллекта – это информационная среда) является основным источником метакогнитивной саморегуляции и самообучения интеллектуальных искусственных систем.

Указываемое исследователями множество ошибок в оценке искусственным интеллектом ситуации и сбоев в его действиях говорит о преждевременном наделении его свойствами метапознания, однако некоторые исследователи утверждают, что именно в разра-

ботке метакогнитивного нейросимволического искусственного интеллекта можно найти решение для оптимизации процесса познания и минимизировать роль человека в обучении [34]. Если ранее языковые модели обучались с подкреплением на основе обратной связи с человеком (RLHF), то сегодня прорывным подходом расширения границ их познавательных возможностей используется обучение с подкреплением на основе обратной связи искусственного интеллекта, с «масштабируемым надзором» (RLAIF) [35]. Возможно ли чисто машинное метапознание и качественный скачок систем искусственного интеллекта к стратегическому метакогнитивному мышлению – вопрос открытый.

Серьёзные препятствия в развитии метапознания у ИИ могут быть обусловлены непредсказуемыми и новыми условиями (надёжность), отсутствием прозрачности в их рассуждениях (объяснимость), трудностями в общении и взаимодействии (сотрудничество) и рисками, связанными с потенциальными вредными действиями (безопасность). Эти факторы объясняются исследователями одним общим недостатком систем ИИ, связанным с «метакогнитивной близорукостью»: искусственному интеллекту не хватает мудрости [36]. Ведь если метапознание напрямую связано с мудростью, то системам ИИ необходимо нарабатывать способность решать неоднозначные, радикально неопределённые, ценностно неоднородные человекообразные задачи, адаптироваться к быстро меняющимся и непредсказуемым условиям и т.д. Но в первую очередь мудростью и совершенством метакогнитивных стратегий должны обладать и разработчики искусственных интеллектуальных систем. Поэтому, на наш взгляд, приоритет в исследованиях естественного и искусственного интеллекта на сегодняшний день должен отдаваться метапознанию в его сопряженности с мудрым использованием жизненного опыта и творчества при производстве знания.

В заключение отметим, что в любом познании как процессе возникает метауровень, который проявляется метапознанием и приводит к выбору или разработке своего метода познания или формированию индивидуальной метакогнитивной стратегии. Метапознание охватывает весь цикл познания, начинающегося от осознания своего незнания и познавательных способностей до формирования завершённого образа знания как элемента сознания. Метакогнитивный стратегический тип мышления достигается формированием и характеризуется определённой степенью зрелости сознания – мудростью.

Перспективы, открывающиеся в области исследования метапознания как для естественного, так и для искусственного интеллекта, отображают качественный переход в исследовательской и творческой работе с такой синтетической сложной системой, как «естественный – искусственный интеллект». Не менее перспективной в университетской среде видится разработка курса по личной эпистемологии или по метапознанию, что позволит конкретизировать и непосредственно применить наработки данных направлений в образовательной практике. Формирование личной эпистемологии в виде познавательных установок, убеждений и приемов познания соответствует обретению собственной метапознавательной позиции, которую и можно назвать «знающим знанием» в фихтовском понимании.

Список источников

1. Карпов А.В., Карпов А.А., Субботина Л.Ю. Методологические и методические основы исследования метакогнитивных детерминант организации деятельности // Российский психологический журнал. 2017. № 1. С. 149–175.
2. Эпистемология вчера и сегодня / отв. ред. В.А. Лекторский. М.: ИФРАН, 2010. 188 с.
3. Flavell J.H. Metacognition and cognitive monitoring a new area of cognitive developmental Inquiry // American Psychologist. 1979. № 34 (10). P. 906–911.
4. Kompa J.S. Metacognition (Part 2): What Makes Us Truly Human? A Literature Review. 2017. URL: <https://joanakompa.com/2017/08/28/metacognition-part-2-what-makes-us-truly-human-a-literature-review/> (дата обращения: 08.09.2024).
5. Халин С.М. Познание и метапознание: проблема типологического единства. Тюмень: Изд-во Тюменского гос. ун-та, 2018. 352 с.
6. Тихонов Р.В., Амалайнен А.В., Морошкина Н.В. Многообразие метакогнитивных чувств: разные феномены или разные термины? // Вестник Санкт-Петербургского университета. Психология и педагогика. 2018. Т. 8, вып. 3. С. 214–242. doi: 10.21638/11701/spbu16.2018.302
7. Russel B. An Inquiry into Meaning and Truth: The William James Lectures for 1940 Delivered at Harvard University. London, N-W: Psychology Press, 1995. 352 p.
8. Чалмерс Д. Сознательный ум: В поисках фундаментальной теории / пер. с англ. М.: УРСС: Книжный дом ЛИБРОКОМ, 2013. 509 с.
9. Zohar A., Ben David A. Paving a clear path in a thick forest: a conceptual analysis of a metacognitive component // Metacognition Learning. 2009. № 4. P. 177–195. doi: 10.1007/s11409-009-9044-6
10. Dinsmore D.L., Alexander P.A., Louglin S.M. Focusing the conceptual lens on metacognition, self-regulation, and self-regulated learning // Educational Psychology Review. 2008. doi: 10.1007/s10648-008-9083-6
11. Фуко М. Слова и вещи. Археология гуманитарных наук. М., 1977. 407 с.
12. Лопарева Т.А. Познательные стратегии и их роль в процессе обучения грамматической стороне говорения // Теория и практика образования в современном мире: материалы III Междунар. науч. конф. (г. Санкт-Петербург, май 2013 г.). СПб.: Реноме, 2013. С. 144–146.
13. Weinstein C., Mayer R. The Teaching of Learning Strategies // Wittrock M. (Ed.) Handbook of Research on Teaching. New York: Macmillan, 1986. P. 315–327.
14. Keestra M. Metacognition and Reflection by Interdisciplinary Experts: Insights from Cognitive Science and Philosophy // Issues in Interdisciplinary Studies. 2017. Vol. 35. P. 121–169.
15. Burkhard R.A. Knowledge Visualization: The Use of Complementary Visual Representations for the Transfer of Knowledge. A Model, a Framework, and Four New Approaches: Dissertation for the degree of Doctor of Sciences. 2005. 156 p. doi: 10.3929/ethz-a-005004486
16. Костромина С.Н., Гнедых Д.С. Информация и знание: подходы к пониманию процессов усвоения информации и формированию знаний в обучении // Вестник ЛГУ им. А.С. Пушкина. 2015. № 2. С. 5–14.
17. Никифоров А.Л. Анализ понятия «знание»: подходы и проблемы // Epistemology & Philosophy of Science. 2009. № 3. С. 61–73.
18. Пигров К.С. Проблема единства знания и технологическая цивилизация // Естественнонаучное и социогуманитарное знание: методологические аспекты взаимодействия: межвуз. сборник. 1990. С. 126–134.
19. Ackoff R.L. From data to wisdom // Journal of Applied Systems Analysis. 1989. № 6. P. 3–9.
20. Басина Н.И., Ковалёва Е.А. Мудрость как социально-философская проблема // Advanced Engineering Research (Rostov-on-Don). 2012. № 2-2 (63). С. 104–109.
21. Merkebu J., Kitsantas A., Durning S.J., Ma T. What is metacognitive reflection? The moderating role of metacognition on emotional regulation and reflection // Frontiers in Education. 2023. Vol. 8. doi: 10.3389/educ.2023.1166195
22. Карпов А.В., Карпов А.А. Структура метакогнитивной регуляции информационной деятельности. Ярославль: Филигрань, 2022. 816 с.
23. Розов М.А. Философия науки в новом видении. М.: Новый хронограф, 2012. 440 с.
24. Агацци Э. Эпистемология и социальное: петля обратной связи // Вопросы философии. 2010. № 7. С. 58–66.
25. Кутырев В.А. Трансцендентальная феноменология как философско-историческое (пред)шествование инфокогнитивной реальности // Философская мысль. 2013. № 7. С. 100–135.
26. Мамардашвили М.К. Стрела познания (набросок естественноисторической гносеологии). М.: Языки русской культуры, 1997. 304 с.
27. Луман Н. Введение в системную теорию. М.: Логос, 2007. 360 с.
28. Аршинов В. И., Буданов В. Г. Системы и сети в контексте парадигмы сложности // Вопросы философии. 2017. № 1. С. 50–61.
29. Fagnant A., Crahay M. Theories of Mind and Personal Epistemology: Their Interrelation and Connection with the Concept of Metacognition // European Journal of Psychology of Education. 2011. Vol. 26, № 2. P. 257–271.
30. Hofer B.K., Bendixen L.D. Personal epistemology: Theory, research, and future directions // Harris K.R., Graham S., Urdan T., McCormick C.B., Sinatra G.M., Sweller J. (Eds.) APA educational psychology handbook. Vol. 1: Theories, constructs, and critical issues. American Psychological Association. 2012. P. 227–256. doi: 10.1037/13273-009
31. Филипенко С.А. Неявное знание в цифровой гуманитаристике // Вопросы философии. 2023. № 9. С. 103–113.
32. Marcus G., Leivada E., Murphy E. A Sentence is Worth a Thousand Pictures: Can Large Language Models Understand Human Language? 2023. URL: <https://arxiv.org/pdf/2308.00109>
33. Ball D.W. The Era of Predictive AI Is Almost Over // The New Atlantis. 2024. 3 77. P. 26–35. URL: <https://www.thenewatlantis.com/publications/the-era-of-predictive-ai-is-almost-over> (дата обращения: 08.09.2024).
34. Wei H. et al. Metacognitive AI: Framework and the Case for a Neurosymbolic Approach // International Conference on Neural-Symbolic Learning and Reasoning. Cham: Springer Nature Switzerland, 2024. P. 60–67.
35. Bay Y. et al. Constitutional AI: Harmlessness from AI Feedback // arXiv preprint arXiv:2212.08073. 2022. URL: <https://arxiv.org/abs/2212.08073> (дата обращения: 21.09.2024).
36. Johnson S.G.B., Karimi A.-H., Bengio Y., Chater N., Gerstenberg T., Larson K., Levine S., Mitchell M., Rahwan I., Schölkopf B., Grossmann I. Imagining and building wise machines: The centrality of AI metacognition // arXiv preprint arXiv:2411.02478. 2024. URL: <https://arxiv.org/abs/2411.02478> (дата обращения: 08.11.2024).

References

1. Karpov, A.V., Karpov, A.A. & Subbotina, L.Yu. (2017) Metodologicheskie i metodicheskie osnovy issledovaniya metakognitivnykh determinant organizatsii deyatel'nosti [Methodological and methodical foundations for the study of metacognitive determinants of activity organization]. *Rossiyskiy psikhologicheskii zhurnal*. 1. pp. 149–175.
2. Lektorskiy, V.A. (ed.) (2010) *Epistemologiya vchera i segodnya* [Epistemology Yesterday and Today]. Moscow: IFRAN.
3. Flavell, J.H. (1979) Metacognition and cognitive monitoring a new area of cognitive developmental Inquiry. *American Psychologist*. 34 (10). pp. 906–911.
4. Kompa, J.S. (2017) *Metacognition (Part 2): What Makes Us Truly Human? A Literature Review*. [Online] Available from: <https://joanakompa.com/2017/08/28/metacognition-part-2-what-makes-us-truly-human-a-literature-review/> (Accessed: 08.09.2024).
5. Khalin, S.M. (2018) *Poznanie i metapoznanie: problema tipologicheskogo edinstva* [Cognition and Metacognition: The Problem of Typological Unity]. Tyumen: University of Tyumen.

6. Tikhonov, R.V., Ammalainen, A.V. & Moroshkina, N.V. (2018) Mnogoobrazie metakognitivnykh chuvstv: raznye fenomeny ili raznye terminy? [The variety of metacognitive feelings: different phenomena or different terms?]. *Vestnik Sankt-Peterburgskogo universiteta. Psikhologiya i pedagogika*. 8 (3). pp. 214–242. doi: 10.21638/11701/spbu16.2018.302
7. Russel, B. (1995) *An Inquiry into Meaning and Truth: The William James Lectures for 1940 Delivered at Harvard University*. London, N-W: Psychology Press.
8. Chalmers, D. (2013) *Soznyayushchii um: V poiskakh fundamentalnoy teorii* [The Conscious Mind: In Search of a Fundamental Theory]. Translated from English. Moscow: URSS: Knizhnyy dom LIBROKOM.
9. Zohar, A. & Ben David, A. (2009) Paving a clear path in a thick forest: a conceptual analysis of a metacognitive component. *Metacognition Learning*. 4. pp. 177–195. doi: 10.1007/s11409-009-9044-6
10. Dinsmore, D.L., Alexander, P.A. & Loughlin, S.M. (2008) Focusing the conceptual lens on metacognition, self-regulation, and self-regulated learning. *Educational Psychology Review*. doi: 10.1007/s10648-008-9083-6
11. Foucault, M. (1977) *Slova i veshchi. Arkheologiya gumanitarnykh nauk* [The Order of Things: An Archaeology of the Human Sciences]. Translated from French. Moscow.
12. Lopareva, T.A. (2013) [Cognitive strategies and their role in teaching the grammatical aspect of speaking]. *Teoriya i praktika obrazovaniya v sovremennom mire* [Theory and Practice of Education in the Modern World]. Proceedings of the III International Conference. St. Petersburg: Renome. pp. 144–146. (In Russian).
13. Weinstein, C. & Mayer, R. (1986) The Teaching of Learning Strategies. In: Wittrock, M. (ed.) *Handbook of Research on Teaching*. New York: Macmillan. pp. 315–327.
14. Keestra, M. (2017) Metacognition and Reflection by Interdisciplinary Experts: Insights from Cognitive Science and Philosophy. *Issues in Interdisciplinary Studies*. 35. pp. 121–169.
15. Burkhard, R.A. (2005) *Knowledge Visualization: The Use of Complementary Visual Representations for the Transfer of Knowledge. A Model, a Framework, and Four New Approaches*. Dissertation for the degree of Doctor of Sciences. doi: 10.3929/ethz-a-005004486
16. Kostromina, S.N. & Gnedykh, D.S. (2015) Informatsiya i znanie: podkhody k ponimaniyu protsessov usvoeniya informatsii i formirovaniyu znaniy v obucheni [Information and knowledge: approaches to understanding the processes of information assimilation and knowledge formation in learning]. *Vestnik LGU im. A.S. Pushkina*. 2. pp. 5–14.
17. Nikiforov, A.L. (2009) Analiz ponyatiya "znanie": podkhody i problemy [Analysis of the concept of "knowledge": approaches and problems]. *Epistemology & Philosophy of Science*. 3. pp. 61–73.
18. Pigrov, K.S. (1990) Problema edinstva znaniya i tekhnologicheskaya tsivilizatsiya [The problem of the unity of knowledge and technological civilization]. In: *Estestvennonauchnoe i sotsiogumanitarnoe znanie: metodologicheskie aspekty vzaimodeystviya* [Natural Science and Socio-Humanitarian Knowledge: Methodological Aspects of Interaction]. Interuniversity collection. S.I. pp. 126–134.
19. Ackoff, R.L. (1989) From data to wisdom. *Journal of Applied Systems Analysis*. 6. pp. 3–9.
20. Basina, N.I. & Kovalyova, E.A. (2012) Mudrost' kak sotsialno-filosofskaya problema [Wisdom as a socio-philosophical problem]. *Advanced Engineering Research* (Rostov-on-Don). 2-2 (63). pp. 104–109.
21. Merkebu, J., Kitsantas, A., Durning, S.J. & Ma, T. (2023) What is metacognitive reflection? The moderating role of metacognition on emotional regulation and reflection. *Frontiers in Education*. 8. doi: 10.3389/educ.2023.1166195
22. Karpov, A.V. & Karpov, A.A. (2022) *Struktura metakognitivnoy regulatsii informatsionnoy deyatel'nosti* [The Structure of Metacognitive Regulation of Information Activity]. Yaroslavl: Filigran.
23. Rozov, M.A. (2012) *Filosofiya nauki v novom videnii* [Philosophy of Science in a New Vision]. Moscow: Novyy khronograf.
24. Agazzi, E. (2010) Epistemologiya i sotsialnoe: petlya obratnoy svyazi [Epistemology and the social: a feedback loop]. *Voprosy filosofii*. 7. pp. 58–66.
25. Kut'yev, V.A. (2013) Transsentsental'naya fenomenologiya kak filosofsko-istoricheskoe (pred'shestvie infokognitivnoy realnosti [Transcendental phenomenology as a philosophical-historical (pre)sequence of infocognitive reality]. *Filosofskaya mysl'*. 7. pp. 100–135.
26. Mamardashvili, M.K. (1997) *Strela poznaniya (nabrosok estestvennoistoricheskoy gnoseologii)* [The Arrow of Cognition (A Sketch of a Natural-Historical Epistemology)]. Moscow: Yazyki russkoy kultury.
27. Luhmann, N. (2007) *Vvedenie v sistemnyuyu teoriyu* [Introduction to Systems Theory]. Moscow: Logos.
28. Arshinov, V.I. & Budanov, V.G. (2017) Sistemy i seti v kontekste paradigmy slozhnosti [Systems and networks in the context of the complexity paradigm]. *Voprosy filosofii*. 1. pp. 50–61.
29. Fagnant, A. & Crahay, M. (2011) Theories of Mind and Personal Epistemology: Their Interrelation and Connection with the Concept of Metacognition. *European Journal of Psychology of Education*. 26 (2). pp. 257–271.
30. Hofer, B.K. & Bendixen, L.D. (2012) Personal epistemology: Theory, research, and future directions. In: Harris, K.R., Graham, S., Urdan, T., McCormick, C.B., Sinatra, G.M. & Sweller, J. (eds) *APA educational psychology handbook*. Vol. 1: Theories, constructs, and critical issues. American Psychological Association. pp. 227–256. doi: 10.1037/13273-009
31. Filipenok, S.A. (2023) Neyavnoe znanie v tsifrovoy gumanitaristike [Tacit knowledge in digital humanities]. *Voprosy filosofii*. 9. pp. 103–113.
32. Marcus, G., Leivada, E. & Murphy, E. (2023) *A Sentence is Worth a Thousand Pictures: Can Large Language Models Understand Human Language?* [Online] Available from: <https://arxiv.org/pdf/2308.00109> (Accessed: 08.09.2024).
33. Ball, D.W. (2024) The Era of Predictive AI Is Almost Over. *The New Atlantis*. 77. pp. 26–35. [Online] Available from: <https://www.thenewatlantis.com/publications/the-era-of-predictive-ai-is-almost-over> (Accessed: 08.09.2024).
34. Wei, H. et al. (2024) Metacognitive AI: Framework and the Case for a Neurosymbolic Approach. *International Conference on Neural-Symbolic Learning and Reasoning*. Cham: Springer Nature Switzerland. pp. 60–67.
35. Bay, Y. et al. (2022) *Constitutional AI: Harmlessness from AI Feedback*. arXiv preprint arXiv:2212.08073. [Online] Available from: <https://arxiv.org/abs/2212.08073> (Accessed: 21.09.2024).
36. Johnson, S.G.B. et al. (2024) *Imagining and building wise machines: The centrality of AI metacognition*. arXiv preprint arXiv:2411.02478. [Online] Available from: <https://arxiv.org/abs/2411.02478> (Accessed: 08.11.2024).

Информация об авторе:

Логиновская Ю.В. – канд. филос. наук, доцент кафедры философии и методологии науки Национального исследовательского Томского государственного университета (Томск, Россия). E-mail: urucel@yandex.ru

Автор заявляет об отсутствии конфликта интересов.

Information about the author:

Yu.V. Loginovskaya, Cand. Sci. (Philosophy), associate professor, National Research Tomsk State University (Tomsk, Russian Federation). E-mail: urucel@yandex.ru

The author declares no conflicts of interests.

Статья поступила в редакцию 26.12.2024;
одобрена после рецензирования 18.01.2025; принята к публикации 31.03.2025.

The article was submitted 26.12.2024;
approved after reviewing 18.01.2025; accepted for publication 31.03.2025.