



DOI: 10.30727/0235-1188-2022-65-1-10-26

Оригинальная исследовательская статья

Original research article

Эпистемологический анализ социогуманитарной значимости новаций искусственного интеллекта в контексте общего искусственного интеллекта*

Д.И. Дубровский

Институт философии РАН Москва, Россия

Аннотация

В последние годы возникли новые направления развития искусственного интеллекта (ИИ), поставлена задача создания общего искусственного интеллекта (ОИИ), который способен выйти за пределы «узкого» ИИ, обрести высокую степень автономности, самостоятельного решения задач в разных условиях внешней среды и таким образом иметь возможность выполнять функции естественного интеллекта. В связи с этим возникают важные философские и теоретико-методологические вопросы, касающиеся определения и оценки социальной значимости новых достижений ИИ, особенно под углом соотношения их социогуманитарных и технологических аспектов. Необходимо преодолеть сугубо технократический подход, который обычно игнорирует отрицательные последствия цифровизации, возможные риски и угрозы развития ИИ. Этому способствует парадигмальный разрыв между системой понятий, используемых для описания технологий, и системами понятий, специфичных для социогуманитарных описаний и объяснений. Возникает так называемый онтологический парадокс при цифровой трансформации и внедрении ИИ в социальные системы. Онтологический парадокс является также и эпистемологическим парадоксом, поскольку всякое утверждение онтологического типа предполагает его эпистемологическое обоснование. Для преодоления указанного разрыва успешно используется предложенная В.Е. Лепским концепция саморазвивающихся полисубъектных сред – кибернетики третьего порядка. Она позволяет создать концептуальный «мост» между двумя системами понятий, не имеющих между собой прямых

* Работа поддержана Российским научным фондом (РНФ), грант № 21-18-00184 «Социогуманитарные основания критериев оценки инноваций, использующих цифровые технологии и искусственный интеллект».

логических связей. Продуктивным инструментом для этого может служить информационный подход, широко используемый для решения подобных междисциплинарных проблем. Он способен теоретически корректно связать технологические и социогуманитарные описания в единой концептуальной структуре и использоваться для анализа выделенных В.Е. Лепским основных социогуманитарных критериев оценки цифровых новаций. В статье подробно рассматриваются основные положения информационного подхода и его применения для разработки систем ОИИ, социогуманитарной оценки автономного развития ИИ и решения вопросов безопасности.

Ключевые слова: «узкий» искусственный интеллект, социогуманитарные критерии, онтологический парадокс, эпистемологический парадокс, саморазвивающиеся полисубъектные среды, информационный подход, безопасность.

Дубровский Давид Израилевич – доктор философских наук, профессор, главный научный сотрудник сектора теории познания Института философии РАН.

ddi29@mail.ru

<https://orcid.org/0000-0003-4392-2526>

Для цитирования: Дубровский Д.И. Эпистемологический анализ социогуманитарной значимости новаций искусственного интеллекта в контексте общего искусственного интеллекта // Философские науки. 2022. Т. 65. № 1. С. 10–26. DOI: 10.30727/0235-1188-2022-65-1-10-26

An Epistemological Analysis of the Social and Humanitarian Significance of Artificial Intelligence Innovations in Context of Artificial General Intelligence*

D.I. Dubrovsky

Institute of Philosophy, Russian Academy of Sciences, Moscow, Russia

Abstract

Nowadays, new directions for the development of artificial intelligence (AI) have emerged, the task has been set to develop artificial general intelligence (AGI), which is able to go beyond the narrow AI, gain a high degree of autonomy, independently solve problems in different environmental conditions and thus have the ability to perform the functions of natural

* The work was supported by the Russian Science Foundation, grant no. 21-18-00184 “Social and humanitarian foundations for evaluation criteria for innovations based on digital technologies and artificial intelligence.”

intelligence. In this regard, important philosophical, theoretical, and methodological questions arise concerning the definition and evaluation of the social significance of new AI achievements, especially regarding the correlation of their socio-humanitarian and technological aspects. It is necessary to overcome a purely technocratic approach, which usually ignores the negative consequences of digitalization, the possible risks and threats of AI development. These difficulties are conditioned by the paradigm gap between the system of concepts used to describe technologies and the systems of concepts specific to socio-humanitarian descriptions and explanations. The so-called ontological paradox arises during digital transformation and the introduction of AI into social systems. An ontological paradox is also an epistemological paradox, since any statement of an ontological type presupposes its epistemological justification. To overcome this gap, V.E. Lepskiy proposes the conception of self-developing polysubjective environments – cybernetics of the third order. This conception allows to create a conceptual “bridge” between two systems of definitions that do not have direct logical connections between them. A productive tool for this can be the information approach, which is widely used to solve such interdisciplinary problems. This makes possible to theoretically correctly connect technological and socio-humanitarian descriptions in a single conceptual structure and to analyze main socio-humanitarian criteria for evaluating digital innovations. The article discusses in detail the main provisions of the information approach and its application for the development of AGI systems, the socio-humanitarian assessment of AI autonomous development, and the resolution of security issues.

Keywords: narrow artificial intelligence, socio-humanitarian criteria, ontological paradox, epistemological paradox, self-developing polysubjective environments, informational approach, security.

David I. Dubrovsky – D.Sc. in Philosophy, Professor, Chief Research Fellow, Department of Theory of Knowledge, Institute of Philosophy, Russian Academy of Sciences.

ddi29@mail.ru

<https://orcid.org/0000-0003-4392-2526>

For citation: Dubrovsky D.I. (2022) An Epistemological Analysis of the Social and Humanitarian Significance of Artificial Intelligence Innovations in Context of Artificial General Intelligence. *Russian Journal of Philosophical Sciences = Filosofskie nauki*. Vol. 65, no. 1, pp. 10–26.

DOI: 10.30727/0235-1188-2022-65-1-10-26

Введение

Нынешний этап развития ИИ, постановка и реализация задачи создания общего ИИ остро ставит новые философские

и теоретико-методологические вопросы, особенно касающиеся необходимости основательного эпистемологического осмысления в единой концептуальной структуре соотношения технологических и социогуманитарных аспектов этого развития. Здесь наблюдается всё еще недопустимый разрыв.

Техногенная цивилизация является по своей сути потребительской, замкнутой в параноидном круге «еще больше производить, чтобы еще больше потреблять, чтобы еще больше производить», и она подчиняет такому курсу развитие самого ИИ, а вместе с этим навязывает и сугубо технократический подход к его объяснению и осмыслению его социальной значимости. Сейчас крайне важно основательно анализировать и оценивать не только действительные социально значимые результаты цифровизации и развития ИИ, но и тех процессов, которые создают явно негативные последствия социогуманитарного порядка. Крупнейшие IT-корпорации в постоянно возрастающем масштабе продуцируют такие новации, которые направлены на разжигание appetites массового потребителя, на создание изохронных средств манипуляции массовым сознанием в своих экономических и политических интересах.

Онтологический и эпистемологический парадоксы

Критический анализ технократического подхода показал его полную концептуальную несостоятельность и позволил вскрыть слишком частую его ангажированность идеологами либерализма, глобализма, западной потребительской системы ценностей. Но в тоже время такой анализ позволил уточнить реальные теоретические трудности использования социогуманитарных критериев при моделировании и реализации функций ИИ и создать условия для разработки новых подходов.

Среди них наиболее значительной и продуктивной является, на наш взгляд, предложенная В.Е. Лепским концепция саморазвивающихся полисубъектных сред – кибернетики третьего порядка [Лепский 2010; Лепский 2017]. Она служит основой для преодоления указанного парадигмального разрыва, того, что именуется **онтологическим парадоксом** при цифровой трансформации и внедрении новаций ИИ в социальные системы.

Но необходимо учитывать, что онтологический парадокс является одновременно и *эпистемологическим парадоксом*, поскольку всякое обоснованное утверждение онтологического типа предполагает его эпистемологическую рефлексию, т.е. рассмотрение достаточности тех познавательных, понятийных средств, которые при этом используются.

Эпистемологический парадокс выражает концептуальный разрыв между такими двумя системами понятий, которые не имеют между собой прямых логических связей. Одна из них служит описанию и объяснению социогуманитарных явлений (это такие понятия, как смысл, ценность, интенциональность, целеполагание, воля, справедливость, доверие и т.п.), вторая – описанию и объяснению физических свойств, устройства и функционирования технологических систем (это понятия физикалистского типа, такие, как масса, энергия, механические взаимодействия, пространственно-временные отношения и структуры и т.п.). Как же возможно приписывать технологическому агенту этические свойства или способность понимания смысла причинных связей? Подобная ситуация, именуемая в аналитической философии «провалом в объяснении», широко обсуждалась многими ее представителями в связи с попытками решения того, что принято называть трудной проблемой сознания, т.е. проблемой объяснения связи между явлениями субъективной реальности, которым нельзя приписывать физические свойства (массу, энергию, пространственные характеристики), и мозговыми процессами. Как известно, предлагавшиеся редукционистские способы решения этой проблемы оказывались мнимыми, так как в результате все те специфические особенности сознания, такие как субъективная реальность, которые требовали объяснения, фактически элиминировались и «провал в объяснении» оставался в силе. В работах [Дубровский 2020; Дубровский 2015; Dubrovsky 2019] предложен способ нередукционистского решения основных теоретических вопросов трудной проблемы сознания на основе информационного подхода.

Сейчас во многих областях науки и технологий настоятельно требуется теоретически обоснованное преодоление

такого провала, что особенно ярко проявляется в развитии ИИ, в том числе при моделировании и реализации описанных В.Е. Лепским основных социогуманитарных критериев оценки цифровых новаций (продуктивность, безопасность, развитие, удовлетворенность) [Лепский 2018], [Лепский 2020]. Для этого требуется определенный категориальный «мост», позволяющий теоретически корректно объединить в единой концептуальной структуре указанные два типа описания и объяснения, которые логически разобщены.

Информационный подход

Как показывает опыт последних десятилетий, категориальный «мост» создается посредством информационного подхода, который широко используется в различных областях науки для решения сложных междисциплинарных проблем. Он может служить эффективным инструментом для многоплановой реализации концепции В.Е. Лепского о саморазвивающихся полисубъектных средах. Несмотря на отсутствие общепринятой теории информации, понятие информации широко используется практически во всех научных дисциплинах. Это обусловлено тем, что понятие «информация» обладает рядом общепринятых определений. Среди них фундаментальное значение имеют следующие: 1) информация необходимо воплощена в своем физическом носителе, не существует вне и помимо него; 2) информация инвариантна по отношению к физическим свойствам своего носителя, т.е. одна и та же для данной системы информация может иметь разные по своим физическим свойствам носители, т.е. кодироваться по-разному; 3) информация в биологических, социальных и социотехнических системах обладает не только синтаксическим, но также *семантическим* и *прагматическим* аспектами, которые релевантны для описания и объяснения таких свойств социогуманитарных явлений как смысл, ценность, цель, волевая активность, вера, управление. Поскольку эти определения общеприняты (по крайней мере, в биологических, социальных и социотехнических дисциплинах, к которым относится и область ИИ), они могут служить основанием для построения в них требуемых теоретических объяснений.

Здесь принципиальное значение приобретает вопрос о *кодировании и декодировании информации*, а также вопрос о способности системы к *самоорганизации*. Всякий носитель информации представляет собой определенную физическую структуру, которая является **кодowym воплощением** данной информации. Она сложилась в ходе биологической эволюции, а затем и социального развития (включая процессы развития и использования технологий). Наряду с кодовыми структурами, типическими для всего класса биологических и социальных систем, отдельные, единичные представители этих классов формируют свои специфичные кодовые структуры, обусловленные индивидуальным опытом их существования, приспособления, обучения. Во всех случаях информация, как таковая, выражает по сути функциональное **значение** для данной самоорганизующейся системы определенного комплекса физических свойств и физических воздействий, сложившегося в виде соответствующей кодовой структуры. Одни физические комплексы имеют для системы временное значение (например, в виде условного рефлекса, отдельных навыков), другие – для всего периода существования данного индивида (скажем, родной язык), а некоторые – для всей истории существования живых существ (генетический код). Информация выполняет функции программирования действий и управления ими. Но лишь в том случае, когда она актуализована, т.е. **декодирована**. Можно выделить два вида кодов: «естественные» и «чуждые». В случае «естественного» кода информация выступает для системы «прозрачной», так сказать, в «чистом» виде, т.е. декодируется автоматически; сразу «понятна», и готова для реализации функции управления; это имеет место в организме на многих уровнях регуляции жизненных процессов, например при мгновенном понимании человеком значения слов родного языка. В «чуждом» коде информация закрыта, «не понятна» для системы. Чтобы она смогла информацию «понять» и использовать, нужна специальная операция декодирования, расшифровки кода, требующая нередко значительных усилий. Это хорошо видно на примерах решения многих познавательных задач, особенно при расшифровке тайных сообщений или зна-

чения слов забытого языка. Опыт решения такого рода задач представляет значительный эпистемологический интерес для разработки и использования информационного подхода в различных научных дисциплинах (см.: [Сингх 2007]).

Нейронаука и «чтение мозга»

Нейронаука широко использует информационный подход для объяснения связи явлений сознания с мозговыми процессами, их декодирования и для объяснения феномена психической причинности. Сравнительно новое направление нейронауки, именуемое «чтением мозга» (brain-reading), используя методы картирования и визуализации мозговых процессов, исследует и выясняет разнообразные нейродинамические корреляты психических явлений. Оно ставит своей задачей расшифровку их мозговых кодов и достигло уже существенных результатов в области изучения явлений субъективной реальности. Еще более десяти лет тому назад японские исследователи Ё. Мияваки, Ю. Камитани и их сотрудники расшифровали нейродинамические эквиваленты зрительного восприятия статичных черно-белых объектов (отводя сигналы от мозга испытуемых на компьютер, воспроизводили на его экране переживаемые ими в данном интервале зрительные образы) [Miyawaki et al. 2008; Fujiwara, Miyawaki, Kamitani 2009]. При этом они смогли расшифровывать не только образы непосредственно воспринимаемого объекта, но и воспоминание о нем. Через некоторое время они научились расшифровывать цветные динамические образы (из кинофильма).

За последние годы на основе технологии глубоких нервных сетей расшифровка мозговых кодов психических явлений значительно усовершенствовалась. Нейроинформатики из Киотского университета в Японии предложили новый способ непосредственно по данным МРТ визуализировать на экране компьютера изображение, которое видит человек в момент сканирования мозга [Shen et al. 2019]. Впечатляющий результат получен в области прямого перевода из мозга в компьютер (т.е. декодирования) текста из 30–50 предложений на английском языке с помощью отведения сигналов примерно от 250 пунктов

коры мозга испытуемого. При этом средняя частота ошибок в словах не превышала 3% [Makin et al. 2018].

К настоящему времени установлено большое число достаточно четких корреляций между определенными психическими явлениями и соответствующими нервными процессами. Они широко используются для создания новых нейротехнологий в медицине, в различных интерфейсах «мозг – компьютер – машина». Они применяются в робототехнике, где их перспективы могут быть особенно значительными в области создания гибридных человеко-робототехнических систем с прямыми интерфейсами «мозг – интеллектуальный робот». Как известно, давно созданы и успешно развиваются средства, позволяющие парализованному человеку мысленно управлять курсором компьютера, инвалидной коляской и даже экзоскелетом. Сюда же относятся мысленно управляемые протезы конечностей. Установленные нейрокорреляты касаются не только действий, но и довольно сложных явлений субъективной реальности, имеющих прямое отношение к процессам мышления и его операциям, к феноменам внимания, намерений человека, выяснения ложных ответов на задаваемые ему вопросы. Все это позволяет считать теоретически возможным и весьма перспективным создание систем мысленного управления роботом.

Психическая и информационная причинность

Требуется основательное теоретическое осмысление вопроса о *психической причинности*. Очевидно, что моя мысль способна вызывать желаемое движение моей руки. Это простейший пример психической причинности. Но ведь по своим мысленным планам мы совершаем сложные системы действий и достигаем желаемых результатов. Как это объяснить, если мысли нельзя приписывать физические свойства? Достаточное теоретическое объяснение может быть дано на основе информационного подхода, согласно которому информация (в данном интервале) и ее физический носитель суть явления однопричинные, одновременные и постольку находятся в отношении взаимно однозначного соответствия. Моя мысль (содержанием которой является «желание и решение взять на столе стакан с водой»)

есть информация в виде сложного ментального образования, включающего наряду с моим желанием и решением мои образы собственной руки, стола, стакана с водой, ряда деталей наличной обстановки. Этот паттерн информации имеет своим физическим носителем определенную кодовую мозговую нейродинамическую структуру, способную запускать соответствующие моторные функции, реализующие мое желание. Здесь описано то, что именуют **информационной причинностью**. Как видим, она необходимо включает физический компонент. Но она отличается от обычной **физической причинности** тем, что вызываемое следствие тут определяется не сугубо физическими свойствами носителя информации, а именно содержанием информации, так как то же самое следствие может быть вызвано носителем этой же информации, имеющей другие физические свойства – в силу *принципа инвариантности информации по отношению к физическим свойствам ее носителя*. **Психическая причинность есть вид информационной причинности**. Помимо нее существуют другие виды информационной причинности (генетические, соционормативные, технические).

Вряд ли надо доказывать фундаментальное значение информационной причинности в процессах цифровизации, в развитии новых направлений ИИ и робототехники, которые отвечают требованиям социогуманитарной значимости, но в то же время и задачам своевременной диагностики и блокирования тех новаций ИИ, которые несут угрозу нашим жизненно важным интересам. Эти вопросы приобретают особенно высокую актуальность в связи с разработкой ОИИ, ставшей сейчас главным мировым трендом в развитии ИИ (в принятой международной номенклатуре ОИИ обозначается сокращенно *AGI*, от *Artificial General Intelligence*).

Разработка ОИИ происходит в условиях нарастающей конкуренции между крупнейшими научными центрами и специализированными корпорациями, – а в более широком масштабе, – между Россией и такими лидирующими в области ИИ государствами, как США, Япония, страны Западной Европы. Наше отставание, однако, не должно вызывать пессимистического настроения. Надо учитывать, что наши конкуренты – лидеры

в области классических направлений ИИ, главным образом, в области «узкого» ИИ. Для успехов же в области ОИИ необходимы принципиально новые теоретические и методологические подходы, новые технологические решения. В этом отношении у нас в стране ведется успешная работа.

Важнейшим условием успешной разработки ОИИ является преодоление ограниченности классической методологии А. Тьюринга, на которой до сих пор основывалось развитие ИИ и выдающиеся успехи информационных технологий, изменивших мир. Но сейчас мы вступили в стратегически новый этап развития ИИ. Методология А. Тьюринга носит сугубо операциональный, бихевиоральный характер и тем самым устраняет использование понятия сознания для описания и моделирования интеллекта. Этот недостаток не раз отмечался рядом специалистов в области ИИ; в последние годы предпринимаются попытки его устранения. Одним из таких наиболее показательных примеров может служить концепция посттьюринговой методологии, предложенная А.Р. Ефимовым, которая существенно расширяет диапазон функций ИИ в интеллектуальной робототехнике [Ефимов 2020].

Требуемые ресурсы

Для успешной разработки ОИИ требуется тщательный анализ пригодных для этих целей ресурсов, имеющихся в различных научных и философских дисциплинах, особенно в нейронауке, психологии, лингвистике и, конечно, в эпистемологии и феноменологических исследованиях структур субъективной реальности, в парадигме постнеклассической рациональности. Значительные ресурсы содержатся в результатах эпистемологического анализа специфических для естественного интеллекта когнитивных структур, касающихся таких феноменов как внимание, категоризация стимулов и действий, дискретизация непрерывных процессов, обобщение и оценка эмпирических данных, соотношение высказываний от первого и от третьего лица, обучение новому навыку, понимание причинных связей и др., т.е. того, что необходимо для ОИИ. Следует подчеркнуть, что большое значение для указанных целей имеют результаты

феноменологических исследований субъективной реальности, позволяющие выделить и описать динамические структуры и операции целереализующего мыслительного процесса, которые могут служить основанием для моделирования специфических когнитивных архитектур ОИИ. Некоторые вопросы этого плана рассмотрены в моей статье [Дубровский 2021].

Вопрос о ресурсах является широким и многоплановым, требует специального анализа. Но здесь хотелось бы еще раз отметить первостепенное значение нейронаучных исследований сознания для проблематики ОИИ. Поскольку отличительным свойством системы ОИИ выступает ее *автономность*, способность самостоятельного решения задач в разных средах, то для моделирования такой способности самообучения может служить описание двуединой функции зеркальных нейронов, которые одновременно кодируют образ предмета и способ действия с ним, что, по мнению исследователей, лежит в основе быстрого обучения новым навыкам. Для понимания и моделирования механизмов самообучения имеют важное значение многие другие достижения нейронауки, раскрывающие механизмы самоорганизации психических процессов [Анохин 2021; Дубровский 2022].

Продвижение в разработке ОИИ будет означать появление все новых функций ИИ, в том числе таких, о которых мы ранее не подозревали, что связано с развитием у технологического агента способности самообучения и самоорганизации. Это ставит сложные вопросы социогуманитарной оценки достижений ОИИ. Одно дело, когда мы проектируем функции робота, число которых ограничено, и заведомо опасные для нас функции могут блокироваться, другое дело, когда робот сам производит новые непредвиденные функции. Такая ситуация должна быть предметом внимания и специального изучения.

Когда речь идет о социогуманитарных оценках новаций ИИ и определении их критериев, то последние имеют два противоположных плана: в одном описывается и оценивается позитивное значение, в другом – негативное. Трудности возникают в связи с тем, что невозможно выстроить строго однозначную иерархическую структуру социогуманитарных ценностей для

всех жизненных ситуаций, т.е. такую, которая бы позволяла во всех случаях производить строго альтернативный выбор. Это хорошо видно при попытках моделирования этически приемлемых функций интеллектуального агента. Например, такая высокая этическая ценность как следование правде и истине в ряде ситуаций вступает в резкое противоречие с другой столь же высокой этической ценностью как, например, сохранение жизни человеку. В свою очередь, сохранение жизни как высшей ценности ставится под вопрос, когда речь идет о таких высоконравственных поступках, как исполнение воинского долга и в других случаях героического самопожертвования в экстремальных ситуациях.

Однако надо подчеркнуть, что во многих видах ситуаций одна определенная этическая ценность, несмотря на ее относительность, приобретает абсолютное значение, не имеет альтернативы и должна быть неукоснительно реализована. Такие этические решения и такие виды ситуаций допускают четкую классификацию для *определенных видов деятельности*, например, при создании интеллектуальных роботов. Сфера видов их деятельности постоянно расширяется, но она все же ограничена. Поэтому при проектировании робота, предназначенного для данного вида деятельности, можно жестко программировать исполнение одного определенного действия и категорический запрет другого.

Тем не менее проблема моделирования этически санкционированных функций у интеллектуального агента, взятая в ее общем виде, остается нерешенной. Судя по существующим исследованиям в той области, которая именуется «этикой искусственного интеллекта», позитивные результаты могут быть достигнуты с помощью построения *частных моделей* на основе классификатора видов деятельности, что позволит осуществить алгоритмизацию определенных этических функций. При этом применяются различные методы (например, метод обучения с подкреплением, используемый для определения оптимального выбора в ситуации неопределенности). В большинстве концептуальных подходов, характерных для проблем «этики искусственного интеллекта», в центре внимания находятся задачи успешного взаимодействия людей с искусственными

автономными системами, в особенности с продвинутыми интеллектуальными роботами, которых учат или которые уже «научились» выполнять ряд санкционированных оператором этически приемлемых действий.

Однако все это большей частью касается лишь функциональных аспектов поведения искусственного интеллектуального агента, остается вне контекста рассмотрения социогуманитарной значимости развития ИИ и, как правило, вне рассмотрения проблематики ОИИ. Между тем именно вопросы разработки ОИИ, уже достигнутые и ожидаемые результаты, приобретают в современных условиях общественной жизни исключительно высокую актуальность. Все более острыми и неотложными становятся проблемы социогуманитарной безопасности, взятые во всем их масштабе и разнообразии. Это вопросы политической, экономической, военной, информационной безопасности и, соответственно, укрепления и дальнейшего развития соответствующих сфер жизнедеятельности.

Заключение

В настоящее время идет ломка системы информационных связей внутри страны и за рубежом, слом мировой финансовой системы и экономики, разжигание межнациональных и иных конфликтов. В этом контексте надо всемерно укреплять безопасность нашей страны во всех областях, включая информационную безопасность, понимаемой как безопасность на всех уровнях и во всех системах и видах коммуникаций, органически связанных с ИИ и ОИИ, в том числе для целей двойного назначения. Это – задача первостепенной политической и социогуманитарной значимости. В ее решении должны принимать самое активное участие не только профессионалы, работающие в области ИИ и ОИИ, масс-медиа, инженерно-технологических специальностей, но и представители социальных и гуманитарных дисциплин.

ЦИТИРУЕМАЯ ЛИТЕРАТУРА

Анохин 2021 – Анохин К.В. В поисках фундаментальной нейронаучной теории // Журнал высшей нервной деятельности. 2021. Т. 71. № 1. С. 39–71.

Дубровский 2020 – *Дубровский Д.И.* Психические явления и мозг. Философский анализ проблемы в связи актуальными задачами нейрофизиологии, психологии и кибернетики / 2-е изд., доп. – М.: ЛЕНАНД, 2020.

Дубровский 2015 – *Дубровский Д.И.* Проблема «Сознание и мозг»: Теоретическое решение. – М.: Канон+, 2015.

Дубровский 2021 – *Дубровский Д.И.* Задача создания Общего искусственного интеллекта и проблема сознания // Философские науки. 2021. № 64. № 1. С. 13–44.

Дубровский 2022 – *Дубровский Д.И.* Значение нейронаучных исследований сознания для разработки общего искусственного интеллекта (методологические вопросы) // Вопросы философии. 2022. № 2. С. 83–93.

Ефимов 2020 – *Ефимов А.Р.* Посттьюринговая методология: разрушение стены на пути к общему искусственному интеллекту // Интеллект. Инновации. Инвестиции. 2020. № 2. С. 74–80.

Лепский 2010 – *Лепский В.Е.* Рефлексивно-активные среды инновационного развития. – М.: Когито-Центр, 2010.

Лепский 2017 – *Лепский В.Е.* Седьмой социогуманитарный технологический уклад – контуры будущего человечества // Глобальный мир: системные сдвиги, вызовы и контуры будущего: XVII Международные Лихачевские научные чтения, 18–20 мая 2017 г. – СПб.: СПбГУП, 2017. С. 357–360.

Лепский 2018 – *Лепский В.Е.* Социогуманитарные критерии оценки новаций цифровой реальности // Социальное время. 2018. № 4 (16). С. 16–26.

Лепский 2020 – *Лепский В.Е.* Философско-методологические основания оценки социально-психологических последствий внедрения новых технологий // Психологический журнал. 2020. Т. 41. № 4. С. 105–108.

Сингх 2007 – *Сингх С.* Книга кодов. Тайная история кодов и их «взлома». – М.: АСТ, Астрель, 2007.

Dubrovsky 2019 – *Dubrovsky D.I.* “The Hard Problem of Consciousness”. Theoretical Solution of its Main Questions // AIMS Neuroscience. Vol. 6. No. 2. P. 85–103.

Fujiwara, Miyawaki, Kamitani 2009 – *Fujiwara Y., Miyawaki Y., Kamitani Y.* Estimating Image Bases for Visual Image Reconstruction from Human Brain Activity // Advances in Neural Information Processing Systems. Vol. 22. P. 576–584.

Makin, Moses, Chang 2020 – *Makin J.G., Moses D.A., Chang E.F.* Machine Translation of Cortical Activity to Text with an Encoder-Decoder Framework // Nature-Neuroscience Technical Report. 2020. Vol. 23. No. 4. P. 575–582.

Miyawaki et al. 2008 – Miyawaki Y., Uchida H., Yamashita O., Sato M.A., Morito Y., Tanabe H.C., Sadato N., Kamitani, Y. Visual Image Reconstruction from Human Brain Activity Using a Combination of Multiscale Local Image Decoders // *Neuron*. Vol. 60. No. 5. P. 915–929.

Shen et al. 2019 – Shen G., Horikawa T., Majima K., Kamitani Y. Deep Image Reconstruction from Human Brain Activity // *PLoS Computational Biology*. Vol. 15, no. 1, e1006633.

REFERENCES

Anokhin K.V. (2021) In Search of a Fundamental Neuroscientific Theory. *Journal of Higher Nervous Activity*. Vol. 71, no. 1, pp. 39–71 (in Russian).

Dubrovsky D.I. (2015) “Consciousness and Brain” Problem: A Theoretical Solution. Moscow: Kanon+ (in Russian).

Dubrovsky D.I. (2019) “The Hard Problem of Consciousness.” Theoretical Solution of Its Main Questions. *AIMS Neuroscience*. Vol. 6, no. 2, pp. 85–103.

Dubrovsky D.I. (2020) *Mental Phenomena and the Brain. Philosophical Analysis of the Problem in Connection with Actual Problems of Neurophysiology, Psychology and Cybernetics* (2nd ed.). Moscow: LENAND (in Russian).

Dubrovsky D.I. (2021) The Task of Creating General Artificial Intelligence and the Problem of Consciousness. *Russian Journal of Philosophical Sciences = Filosofskie nauki*. Vol. 64, no. 1, pp. 13–44 (in Russian).

Dubrovsky D.I. (2022) The Value of Neuroscience Research of Consciousness for the Development of Artificial General Intelligence (Methodological Issues). *Voprosy filosofii*. No. 2, pp. 83–93 (in Russian).

Efimov A.R. (2020) Post-Turing Methodology: Breaking the Wall toward Artificial General Intelligence. *Intellekt. Innovatsii. Investitsii*. No. 2, pp. 74–80 (in Russian).

Fujiwara Y., Miyawaki Y., & Kamitani Y. Estimating Image Bases for Visual Image Reconstruction from Human Brain Activity. *Advances in Neural Information Processing Systems*. Vol. 22, pp. 576–584.

Lepskiy V.E. (2010) *Reflexively Active Environments of Innovative Development*. Moscow: Kogito-tsentr (in Russian).

Lepskiy V.E. (2017) The Seventh Socio-Humanitarian Technological Order – the Outlines of the Future of Mankind. In: Zapesotsky A.S., Markov A.P., Paseshnikova L.A. (Eds.) *Global World: Systemic Shifts, Challenges and Outlines of the Future: 17th International Likhachev Scientific Readings, May 18–20, 2017* (pp. 357–360). Saint Petersburg: SPbGUP (in Russian).

Lepskiy V.E. (2018) Socio-Humanitarian Criteria for Evaluating the Innovations of Digital Reality. *Sotsial'noye vremya*. No. 4, pp. 16–26 (in Russian).

Lepskiy V.E. (2020) Philosophical and Methodological Foundations for Assessing the Socio-Psychological Consequences of the Introduction of New Technologies. *Psychological Journal*. Vol. 41, no. 4, pp. 105–108 (in Russian).

Makin J.G., Moses D.A., & Chang E.F. (2020) Machine Translation of Cortical Activity to Text with an Encoder-Decoder Framework. *Nature-Neuroscience Technical Report*. Vol. 23, no. 4, pp. 575–582.

Miyawaki Y., Uchida H., Yamashita O., Sato M.A., Morito Y., Tanabe H.C., Sadato N., & Kamitani Y. (2008) Visual Image Reconstruction from Human Brain Activity Using a Combination of Multiscale Local Image Decoders. *Neuron*. Vol. 60, no. 5, pp. 915–929.

Shen G., Horikawa T., Majima K., & Kamitani Y. (2019) Deep Image Reconstruction from Human Brain Activity. *PLoS Computational Biology*. Vol. 15, no. 1, e1006633.

Singh S. (1999) *The Code Book: The Science of Secrecy from Ancient Egypt to Quantum Cryptography*. New York: Doubleday (Russian translation: Moscow: AST, Astrel, 2007).