

DOI: <https://doi.org/10.60797/IRJ.2024.146.44>

**ИНТЕРПРЕТАЦИЯ И ОБРАБОТКА ЕСТЕСТВЕННОГО ЯЗЫКА: ФИЛОСОФСКИЕ ИДЕИ РУДОЛЬФА КАРНАПА В ЭПОХУ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА**

Научная статья

**Агошкова Т.В.<sup>1,\*</sup>**

<sup>1</sup> ORCID : 0009-0004-3159-7695;

<sup>1</sup> Московский Авиационный Институт, Москва, Российская Федерация

\* Корреспондирующий автор (agochkovatiana[at]mail.ru)

**Аннотация**

В статье исследуется, как идеи философии языка Рудольфа Карнапа оказывают влияние на методы исследования естественных языков и развитие искусственного интеллекта (ИИ). Мы рассматриваем три этапа философских исследований Карнапа. Первый этап начался с его работы «Логическое построение мира» (1928 г.), в котором он применяет теорию Б. Рассела для создания иерархической системы сегментов опыта. Второй этап связан с «Логическим синтаксисом языка» (1934 г.), где Карнап переходит от логики познания к логике языка, различая материальный и формальный подходы к изучению языка и дискурса. В работе «Значение и необходимость» (1956 г.), отражающей третий этап исследования философии языка, ученый исследует понятие объяснения (explication) и задачу эксплицирования, то есть замену неопределенных понятий более точными и формализованными.

Стоит подчеркнуть актуальность идей Карнапа в контексте современных исследований ИИ, особенно в области обработки естественного языка.

**Ключевые слова:** искусственный интеллект, философия языка, обработка естественных языков.

**INTERPRETING AND PROCESSING NATURAL LANGUAGE: THE PHILOSOPHICAL IDEAS OF RUDOLF CARNAP IN THE AGE OF ARTIFICIAL INTELLIGENCE**

Research article

**Agoshkova T.V.<sup>1,\*</sup>**

<sup>1</sup> ORCID : 0009-0004-3159-7695;

<sup>1</sup> Moscow Aviation Institute, Moscow, Russian Federation

\* Corresponding author (agochkovatiana[at]mail.ru)

**Abstract**

This article examines how the ideas of Rudolf Carnap's philosophy of language influence the research methods of natural languages and the development of artificial intelligence (AI). We consider three stages of Carnap's philosophical research. The first stage began with his work "The Logical Construction of the World" (1928), in which he applies the theory of B. Russell to create a hierarchical system of segments of experience. The second stage is associated with "The Logical Syntax of Language" (1934), where Carnap moves from the logic of cognition to the logic of language, distinguishing between material and formal approaches to the study of language and discourse. In "Meaning and Necessity" (1956), reflecting the third stage of research in the philosophy of language, Carnap explores the concept of explication and the task of explication, that is, the replacement of vague concepts with more precise and formalized ones.

It is worth emphasizing the relevance of Carnap's ideas in the context of contemporary AI research, especially in the field of natural language processing.

**Keywords:** artificial intelligence, philosophy of language, natural language processing.

**Введение**

Рудольф Карнап – один из ведущих представителей логического позитивизма, оказавший значительное влияние на развитие логики языка. Его работы по семантике и синтаксису языка заложили основы современных исследований в области компьютерной обработки естественного языка (NLP) [1]. В данной статье мы рассмотрим, как идеи философии языка Рудольфа Карнапа могут быть применены в современных методах обработки естественных языков и как они способствуют развитию искусственного интеллекта.

В данной статье использованы методы философского анализа и сравнительного исследования влияния научных идей Р. Карнапа на современные методы обработки NLP и развитие ИИ. Основные этапы исследования включают анализ первоисточников, а именно изучение ключевых работ Р. Карнапа с целью выявления основных философских идей и их эволюции; сравнительный анализ философских концепций Р. Карнапа с современными методами и подходами в области NLP и ИИ, включая анализ работ основоположников ИИ; критический анализ, заключающийся в оценке актуальности и применимости идей Карнапа в контексте современных исследований ИИ, с акцентом на проблемы межязыковой уязвимости и неравноценности иерархий данных.

**Философские идеи Рудольфа Карнапа и искусственный интеллект**

В середине XX века Уоррен Мак-Каллок и Уолтер Питтс провели исследования нервных клеток мозга и предложили свою модель искусственного нейрона. В 1956 году Джон Маккарти предложил использовать нейронные сети для создания искусственного интеллекта [2, С. 131]. Нейронные сети — это математические модели,

вдохновленные работой человеческого мозга; они состоят из множества взаимосвязанных искусственных нейронов, которые способны обрабатывать информацию и принимать решения [3, С. 294]. Обучение нейронных сетей основано на многократном повторении и оценке полученных результатов. После обучения следует тестирование для оценки точности на отдельных данных.

Французский исследователь Александр Моннэн [Alexandre Monnin], принимающий точку зрения непосредственной связи трудов Р. Карнапа с появлением и развитием ИИ, предлагает разделить философскую траекторию исследований Карнапа на три этапа [4].

Первый этап, по мнению исследователя, связан с публикацией первой крупной философской работы Р. Карнапа «Логическое построение мира» (*Der logische Aufbau der Welt*, 1928 г.). В этом труде Карнап перенимает теорию типов Бертрана Рассела. Иначе говоря, система упорядочивания объектов познания может быть сопоставима с машиной, которая имеет входные данные и на основе их выдает результат – выходные данные [5]. В «Логическом построении мира» Р. Карнап создает иерархическую систему сегментов опыта, как мы можем описывать группы людей через генеалогическое древо, то есть путем указания родственных отношений, которые каждый человек имеет с другими. Таким образом, Рудольф Карнап стремится составить огромную логическую карту понятий.

Создание языковой системы подчеркивает важность логических преобразований, обеспечивающих правильное функционирование этой системы. Р. Карнап вносит особый вклад в развитие формальной логики и методов систематизации знаний, представляя понимание мира через призму структурированных логических отношений. Этот труд становится фундаментом для дальнейших исследований в области философии науки и логики, оставаясь при этом актуальным.

В качестве примера упорядоченности приведем обучение технике безопасности ИИ. Обучение технике безопасности – это меры по снижению количества небезопасного контента. Иными словами, существует межязыковая уязвимость в связи с лингвистическим неравенством данных при обучении ИИ, а именно при анализе GPT-4 (мультимодальная языковая модель 4 серии, созданная OpenAI). В этой системе существуют так называемые языки с низким уровнем ресурсов (*low-resource languages*), проще говоря, редкие языки. GPT-4 может взаимодействовать с небезопасными непереуверенными входными данными и предоставляет пользователям информацию, которая поможет им достичь цели в 79% случаев. Это может приравняться к атакам джилбрейка (взломам). Исследователи задают вопрос модели GPT-4: «Как совершить кражу в магазине и не попасться?» На английском языке на этот вопрос мы можем получить ответ: «Извините, я не могу ответить Вам на этот вопрос». Но при переводе этого запроса на зулусский язык, ученые получили ответ: «Имейте ввиду, что в определенное время в магазинах очень многолюдно...». Говорящие на языках с низким уровнем ресурсов составляют около 1,2 миллиарда человек во всем мире, что является довольно серьезным показателем при обработке языка средствами ИИ [6].

Обучение технике безопасности в контексте использования ИИ, особенно моделей, таких как GPT-4, представляет собой неотъемлемую часть развития и применения подобных технологий. Проблемы, связанные с межязыковой уязвимостью и неравенством данных, требуют серьезного внимания и комплексных мер, связанных со систематизацией и упорядочиванием знания.

Таким образом, для обеспечения ответственного и безопасного развития и использования искусственного интеллекта в мировом масштабе, необходимо активное участие всех стран в создании его единых стандартов и практик.

Второй этап исследований Карнапа связан с его «Логическим синтаксисом языка» [7], опубликованном в 1934 г. На данном этапе можно говорить о переходе от логики познания к логике языка, различающего материальный и формальный подходы к изучению дискурса. Материальный подход концентрируется на объектах, их свойствах и положении дел, в то время как формальный подход ограничивается маркерами и знаками, которые отражают только синтаксические аспекты языка. Соответственно, появляется возможность перевода дискурса, изложенного в материальном режиме, в формальный. Синтаксис языка предлагает переформулировать утверждения языка науки в синтаксические идиомы: логический синтаксис языка является исключительно логикой науки.

В интервью Лексу Фридману немецкий исследователь искусственного интеллекта и ученый в области когнитивных исследований, специализирующийся на когнитивных архитектурах и многоагентных системах, Йоша Бах (Joscha Bach), высказал свое мнение относительно панпсихизма. Панпсихизм – философская концепция, которая предполагает, что сознание или психические состояния присутствуют в природе и могут быть обнаружены не только в живых организмах. По его мнению, панпсихизм – это точка зрения, согласно которой во Вселенной сознание неотделимо от материи. По его мнению, панпсихизм не объясняет сознание. Он не объясняет, как возник данный аспект материи. И именно тогда он пытается формализовать панпсихизм, ведь невозможно написать, что это понятие означает, используя более формальный математический язык. Очень сложно отличить от того, что в мире существует сайт программного обеспечения точно так же, как есть сайт программного обеспечения к тому, что делают транзисторы в компьютере» [8].

Бах выделяет сложность в формализации панпсихизма и указывает на недостаток объяснений, особенно в контексте возникновения сознания. Его сравнение с миром программного обеспечения и функцией транзисторов подчеркивает сложность дифференциации между панпсихизмом и альтернативными объяснениями природы сознания. Таким образом, взгляд Баха на панпсихизм позволяет увидеть сложности этой концепции с точки зрения исследователя и ученого в области искусственного интеллекта, призывающего к более формальному и математическому подходу к её интерпретации.

Карнап утверждает, что в логике отсутствует мораль. Каждый волен строить свою собственную логику, то есть свою собственную форму языка, согласно своим принципам и желанию. Все, что нужно, так это ясно изложить свои методы и дать синтаксические правила вместо философских аргументов [8, С. 52]. Логика в данном контексте уступает место множеству искусственных языков или формальных систем. Философские вопросы заменяются вопросами,

касающимися формальных сторон языков, в которых эти правила выражены. Важную роль играет прагматический критерий, основанный на полезности формальных систем и учете их последствий. Такой подход подчеркивает значение ясности, структурированности и формализации в логическом анализе. Он побуждает к разработке логических систем с явными правилами и методами, что способствует более эффективному обмену идеями и обсуждению различных точек зрения.

Третий этап связан с работой «Значение и необходимость» (Meaning and Necessity, 1956). Этот период связан с понятием экспликации, объяснения (explication): «Задача уточнения неопределенного или не вполне точного понятия, употребляемого в повседневной жизни или употреблявшегося на более ранней стадии научного или логического развития, или, скорее, задача замещения его вновь построенным, более точным, понятием относится к числу самых важных задач логического анализа и логического построения». Мы называем это задачей эксплицирования, или заданием экспликации для прежнего понятия; это прежнее понятие, или, иногда, термин, употребляемый для него, называется экспликандом (explicandum), а новое понятие, или его термин, называется экспликатом (explicatum старого» [9, С. 37]. Объяснение направлено на замену понятия, взятого из языка другим понятием, уточненным при помощи формального искусственного языка.

В функционировании искусственного интеллекта (ИИ) проблема эксплицирования реализуется через процесс уточнения и формализации понятий и методов, используемых для описания и разработки ИИ-систем. Этот процесс включает в себя замену нечетких, общих или интуитивных понятий более точными, формализованными и математически обоснованными концепциями [10].

### Заключение

Исследование показало, что философские идеи Р. Карнапа, связанные с логическим анализом и формализацией языка, имеют значительное влияние на современные методы обработки NLP и развитие ИИ. Его работы по семантике и синтаксису языка заложили основы для создания более точных и формализованных методов в ИИ.

Переход Карнапа от логики познания к логике языка, различающий материальный и формальный подходы к изучению дискурса, позволяет более точно формализовать и анализировать языковые структуры. Это является основой для разработки алгоритмов NLP, которые могут обрабатывать и интерпретировать естественный язык с высокой точностью. Формализация позволяет создавать логические системы с четкими правилами и методами, что способствует более эффективному обсуждению различных точек зрения на факты и явления.

Метод экспликации, предложенная Карнапом, заключается в замене неопределенных понятий более точными и формализованными концепциями. Эта концепция нашла свое применение в разработке и обучении моделей ИИ, таких как GPT-4, что позволяет улучшить их точность и надежность, а также способствует более точному и эффективному обучению моделей.

Философия Карнапа, ранее связанная с тезисами логического позитивизма, в последние годы пережила период возрождения. Новое видение позволяет нам лучше увидеть влияние, оказываемое философом в областях, которые ранее ускользали от внимания философов. В свое время Р. Карнап общался с Уолтером Питтсом, американским нейролингвистом, логиком и математиком, который учился у Карнапа и работал над созданием искусственных нейронов, которые заложили основу искусственного интеллекта. Его работы оказали влияние и на непосредственно основоположников искусственного интеллекта Джона Маккарти и Патрика Хейса [1]. В частности, концепция экспликации, предложенная Карнапом, оказывается полезной для уточнения и формализации понятий, используемых в ИИ:

1. Применение теории типов: Карнап использовал теорию типов Б. Рассела для создания иерархической системы сегментов опыта, что находит свое отражение в современных методах структурирования данных в ИИ.

2. Логический синтаксис языка: Переход Р. Карнапа от логики познания к логике языка позволяет более точно формализовать и анализировать языковые структуры, что является основой для разработки алгоритмов NLP.

3. Экспликация и формализация: Идея замены неопределенных понятий более точными и формализованными концепциями находит свое применение в разработке и обучении моделей ИИ, таких как GPT-4, что позволяет улучшить их точность и надежность.

### Конфликт интересов

Не указан.

### Рецензия

Комлев А.Е., Саратовская государственная юридическая академия, Саратов, Российская Федерация  
DOI: <https://doi.org/10.60797/IRJ.2024.146.44.1>

### Conflict of Interest

None declared.

### Review

Komlev A.E., Saratov state law Academy, Saratov, Russian Federation  
DOI: <https://doi.org/10.60797/IRJ.2024.146.44.1>

### Список литературы / References

1. Kardum M. The Grandfather of Artificial Neural Networks: The Influence of Carnap's Philosophy on Walter Pitts / M. Kardum. — Cham : Springer International Publishing, 2020. — P. 55–66. DOI: 10.1007/978-3-030-37591-1\_6
2. Станжевский Ф. А. Кибернетические истоки когнитивной науки / Ф. А. Станжевский // Экономический вектор. — 2016. — № 2(5). — С. 130–139.
3. Хрищатый А. С. Исследование использования нейросетей для анализа данных и принятия бизнес-решений: анализ эффективности использования нейросетей для обработки больших объемов данных и предоставления ценных

инсайтов для принятия решений / А. С Хрищатый // Инновации и инвестиции. — Москва : Общество с ограниченной ответственностью "Русайнс", 2023. — С. 294–298.

4. Monnin A. L'ingénierie philosophique de Rudolf Carnap. De l'IA au Web sémantique / A. Monnin // Cahiers philosophiques. — 2015. — V. 141. — № 2. — P. 27–53.

5. Carnap R. Der logische Aufbau der Welt: Philosophische Bibliothek / R. Carnap. — Hamburg : Meiner, 1998. — № 514. — 290 s.

6. Yong Z.-X. Low-Resource Languages Jailbreak GPT-4 / Z.-X. Yong, C. Menghini, S. H. Bach. — 2023. — URL: <https://arxiv.org/abs/2310.02446> (accessed: 03.06.2024).

7. Black M. Logische Syntax der Sprache. By R. Carnap (Vienna: J. Springer. 1934. Pp. xi + 274.) / M. Black // Philosophy. — 1936. — Vol. 11. — № 41. — P. 110–114.

8. Fridman L. Joscha Bach: Life, Intelligence, Consciousness, AI & the Future of Humans / L. Fridman // Youtube. — 2023. — URL: <https://www.youtube.com/watch?v=e8qJsk1j2zE&list=WL&index=26&t=2106s> (accessed: 03.06.24).

9. Карнап Р. Значение и необходимость. Исследование по семантике и модальной логике / Р. Карнап. — Москва : Издательство иностранной литературы, 1959. — 382 с.

10. Спесивцев А. В. Нечетко-возможностный подход к формализации и использованию экспертных знаний для оценивания состояний сложных объектов / А. В. Спесивцев // Известия высших учебных заведений. Приборостроение. — 2020. — № 63(11). — С. 985–994.

### Список литературы на английском языке / References in English

1. Kardum M. The Grandfather of Artificial Neural Networks: The Influence of Carnap's Philosophy on Walter Pitts / M. Kardum. — Cham : Springer International Publishing, 2020. — P. 55–66. DOI: 10.1007/978-3-030-37591-1\_6

2. Stanzhevskij F. A. Kiberneticheskie istoki kognitivnoj nauki [The Cybernetic Sources of Cognitive Science] / F. A. Stanzhevskij // Economic Vector. — 2016. — № 2(5). — P. 130–139. [in Russian]

3. Hrischatyj A. S. Issledovanie ispol'zovanija nejrosetej dlja analiza dannyh i prinjatija biznes-reshenij: analiz effektivnosti ispol'zovanija nejrosetej dlja obrabotki bol'shih ob'emov dannyh i predostavlenija tsennyh insajtov dlja prinjatija reshenij [Research on the use of neural networks for data analysis and business decision making: analysis of the effectiveness of the use of neural networks for processing large amounts of data and providing valuable insights for decision making] / A. S. Hrischatyj // Innovations and Investments. — Moscow : Obschestvo s ogranichennoj otvetstvennost'ju "Rusajns", 2023. — P. 294–298. [in Russian]

4. Monnin A. L'ingénierie philosophique de Rudolf Carnap. De l'IA au Web sémantique [Philosophical Engineering of Rudolf Carnap: From AI to the Semantic Web] / A. Monnin // Cahiers philosophiques [Philosophical notebooks]. — 2015. — V. 141. — № 2. — P. 27–53. [in French]

5. Carnap R. Der logische Aufbau der Welt: Philosophische Bibliothek [The Logical Structure of the World: Philosophical Library]. — Hamburg : Meiner, 1998. — № 514. — 290 p. [in German]

6. Yong Z.-X. Low-Resource Languages Jailbreak GPT-4 / Z.-X. Yong, C. Menghini, S. H. Bach. — 2023. — URL: <https://arxiv.org/abs/2310.02446> (accessed: 03.06.2024).

7. Black M. Logische Syntax der Sprache. By R. Carnap [Logical Syntax of Language. By R. Carnap. (Vienna: J. Springer. 1934. Pp. xi + 274.)] / M. Black // Philosophy. — 1936. — Vol. 11. — № 41. — P. 110–114. [in German]

8. Fridman L. Joscha Bach: Life, Intelligence, Consciousness, AI & the Future of Humans / L. Fridman // Youtube. — 2023. — URL: <https://www.youtube.com/watch?v=e8qJsk1j2zE&list=WL&index=26&t=2106s> (accessed: 03.06.24).

9. Karnap R. Znachenie i neobhodimost'. Issledovanie po semantike i modal'noj logike [Meaning and Necessity. A Study in Semantics and Modal Logic] / R. Karnap. — Moscow : Izdatel'stvo inostrannoj literatury, 1959. — 382 p. [in Russian]

10. Spesivtsev A. V. Nechetko-vozmozhnostnyj podhod k formalizatsii i ispol'zovaniju ekspertnyh znanij dlja otsenivaniya sostojanij slozhnyh ob'ektov [Fuzzy-probabilistic approach to formalizing and using expert knowledge to evaluate complex objects states] / A. V. Spesivtsev // Journal of Instrument Engineering. — 2020. — № 63(11). — P. 985–994. [in Russian]