



НАУЧНЫЕ ОТКРЫТИЯ МЕТОДАМИ LLM: ФАКТЫ И ИНТЕРПРЕТАЦИЯ

SCIENTIFIC DISCOVERIES BY LLM METHODS: FACTS AND INTERPRETATION

В. Ю. Филимонов*

ORCID: 0009-0000-6795-507X

Vladimir Yu. Filimonov*

* *Пятигорский государственный университет, Пятигорск, Россия*

* *Pyatigorsk State University, Pyatigorsk, Russia*

Цель исследования — проведение анализа уровней проявления категорий факта и интерпретации в больших языковых моделях.

Методологическую базу исследования составили разработки категорий модели социального конструктивизма Т. Сейновски, модель сенсорного мышления искусственного интеллекта Д. Дж. Чалмерса, модель этических паттернов искусственного интеллекта В. Н. Соколыч, модель интерпретации текста в системе больших языковых моделей Б. В. Орехова.

Результаты исследования. Проведенное исследование позволило систематизировать проявления свойств факта и интерпретации в следующих представлениях LLM: как интерпретатора мнения личности, как интерпретатора возможной межличностной дискуссии, а также как интерпретатора научных фактов и тематических исследований в рамках информационных баз данных. Определены следующие проблемные направления развития LLM как интерпретатора и концептуальной основы развития области научных открытий: трудности в генерации LLM-контента для редких или малоизученных явлений; некорректные и фальсифицированные

Objective of the study. The analysis of the levels of manifestation of such categories as “fact” and “interpretation” in large language models.

Methodological basis of the study includes the developments of the categories of the social constructivism model by T. Sejnowski, the model of sensory thinking of artificial intelligence by D. J. Chalmers, the model of ethical patterns of artificial intelligence by V. N. Sokolchik, the model of text interpretation in the system of large language models by B. V. Orekhov.

The results of the study. The conducted research allows us to systematize the manifestations of the properties of fact and interpretation in the following representations of LLM: as an interpreter of personal opinion, as an interpreter of possible interpersonal discussion, as well as an interpreter of scientific facts and case studies in the framework of information databases. The following problematic areas of LLM development as an interpreter and a conceptual framework for the development of the field of scientific discoveries are identified: difficulties in generating LLM content for rare or poorly studied phenomena; incorrect and falsified bibliographic references; LLM errors in elementary

библиографические ссылки; ошибки LLM в элементарных заданиях; самостоятельное изобретение LLM новых слов, не свойственных носителям языка.

Перспективы исследования. В рамках исследования сформированы подходы к развитию метафоры о возможности представления факта и интерпретации в виде средств оценки информационных данных больших языковых моделей в различных субъективных формах: от личностной интерпретации к интерпретациям различных массивов крупных научных данных.

Ключевые слова: факт, интерпретация, научные открытия, большие языковые модели, LLM, методы LLM

tasks; independent invention of new words by LLM that are not typical for native speakers.

Prospects of the study. Within the framework of the research, the author forms the approaches to the development of a metaphor about the possibility of representing fact and interpretation in the form of means of evaluating information data from large language models in various subjective forms: from personal interpretation to interpretations of various arrays of scientific data.

Keywords: fact, interpretation, scientific discoveries, large language models, LLM, LLM methods

Введение

Процесс получения научных результатов с целью совершения научных открытий в различных областях — актуальная цель науки не только в ретроспективе, но и в ее современном виде. Развитие инструментов информационного анализа, которые помогают анализировать крупные объемы текстовых данных, а затем их синтезировать, помогает постепенно приближаться к возможности получения генеративных научных данных, а впоследствии — и генеративных научных открытий. В результате такие составные категории научных открытий, как «факт» и «интерпретация», становятся не только инструментом ручного анализа той или иной информации, но и инструментом информационного анализа внутри больших языковых моделей.

Уровни проявления факта и интерпретации в больших языковых моделях

Рассмотрим термины «факт», «интерпретация» и «большая языковая модель» (LLM).

Факт (от лат. *factum* — сделанное, совершившееся) в философии науки — особого рода предложения, фиксирующие эмпирическое знание¹.

¹ Факт. Новая философская энциклопедия Института Философии Российской Академии Наук. Институт Философии Российской Академии Наук. — URL: <https://iphras.ru/elib/3147.html>. (дата обращения: 26.07.2025).

Интерпретация (от лат. *interpretatio* — разъяснение, истолкование) — в широком смысле слова истолкование, объяснение какой-либо реальной ситуации или идейной позиции; в качестве специального понятия методологии науки, опирающейся на семиотический анализ языка науки, интерпретация означает процедуру придания смысла формальным конструкциям языка науки, в результате которой последние превращаются в содержательные термины или утверждения¹.

Большая языковая модель (LLM) — это языковая модель, обученная с помощью машинного обучения без учителя на большом объеме текста и предназначенная для решения задач обработки естественного языка, особенно генерации текста².

Поскольку факт и интерпретация способны постепенно переходить в рамки информационной деятельности больших языковых моделей, определим следующие уровни их проявления в данных моделях:

1. LLM как интерпретатор мнения личности

Т. Сейновски, стоя на позициях социального конструктивизма, рассуждает о «социальной нейросоциологии», где искусственные нейросети взаимодействуют с социальными сетями (Сейновски, 2022, С. 25). То есть социальные сети начинают рассматриваться как полноценная платформа для анализа информационных потоков, где базисом служит коммуникация людей на открытой информационной основе — между людьми и между сетевыми сообществами. Фактологические данные свидетельствуют о том, что за полгода использования GPT-4 когнитивные навыки данной нейросети существенным образом регрессировали за счет необходимости работы с несуразными вопросами. Можно свидетельствовать о том, что произошла трансформация нейросети из «супермозга в двоечника-подростка». А может быть, нейросеть притворяется глупой, пообщавшись с другими нейросетями? (Кузьминов, 2024). То есть, в субъектном плане, нейросеть с базисом большой языковой модели может условным образом ассоциироваться с когнитивным статусом объекта интерпретации, в качестве которого выступает та или иная личность. Однако что если этой личностью является известный медийный деятель со значимыми когнитивными способностями?

К примеру, новейшая модель искусственного интеллекта от xAI Grok 4 при ответе на вопросы об израильско-палестинском конфликте, абортах и иммиграционном законодательстве ссылается на посты в социальных сетях

¹ Интерпретация. Новая философская энциклопедия Института Философии Российской Академии Наук. Институт Философии Российской Академии Наук. — URL: <https://iphras.ru/elib/3147.html>. (дата обращения: 26.07.2025).

² Large language model. Wikipedia. — URL: https://en.wikipedia.org/wiki/Large_language_model. (дата обращения: 26.07.2025).

с аккаунта Илона Маска. Grok также, похоже, ссылается на позицию Маска по спорным вопросам в новостных статьях¹. В данном случае LLM, как интерпретатор, получает более выраженные управленческие возможности по формированию того или иного научно-образовательного дискурса, где за нейросетью стоит мнение конкретного человека, пользующегося во многом выраженной социальной поддержкой.

В обоих рассмотренных случаях детерминанты «факта» и «интерпретации» размываются, поскольку личностная позиция при их формировании особенно субъективна с точки зрения как эмоционального, так и управленческого подкрепления. Если рассматривать мнение личности как потенциал для развития научных открытий методами LLM, то необходимо использовать методы анализа, которые позволят конкретизировать проявления факта и интерпретации, а также избавиться от их сугубо субъективной трактовки.

Чалмерс склоняется к следующему выводу: «Чистые БЯМ (LLM) (например, мультимодальные модели) могут обладать сенсорным опытом и полноценными сенсорными концепциями, по крайней мере, когда они обрабатывают соответствующие текстовые переводы» (Chalmers, 2023).

Не вдаваясь в подробное критическое рассмотрение аргументации Чалмерса, можно отметить следующее: в трактовке БЯМ как чистого говорящего присутствует прямое указание на способность понимать естественный язык. Если это не просто игра слов, то, как показывает обсуждение в рамках работ Д. В. Зайцева, именно этой способности и недостает языковым моделям (Зайцев, 2024).

Если принять во внимание тот факт, что главным этическим паттерном взаимодействия с ИИ в процессе публикации является полная ответственность автора (человека) как за процесс получения знания, так и за его результат, включая возможные ошибки, фальсификации, предубеждения и т. д., сгенерированные ИИ (Соколичик, 2024), то можно отметить, что ответственность за распространение мнения личности через его интерпретацию посредством LLM смещается в зону ответственности субъекта-пользователя таким интерпретированным мнением.

2. LLM как интерпретатор возможной межличностной дискуссии

И. О. Вагин, российский бизнес-тренер, поделился результатом работы с нейросетью ChatGPT по генерации возможной дискуссии политических лидеров нашей истории: Сталина, Горбачева, Ельцина и Ленина на тему

¹ Grok 4 seems to consult Elon Musk to answer controversial questions. Techcrunch. — URL: <https://techcrunch.com/2025/07/10/grok-4-seems-to-consult-elon-musk-to-answer-controversial-questions/> (дата обращения: 28.07.2025).

«Современная Россия»¹. В рамках приведенной беседы сохраняется дипломатический дискурс, при этом мнение отдельной личности вступает в межличностное взаимодействие с другими мнениями с целью выработки оптимальных и логически связанных положений по тематике беседы. В то же время можно выявить некоторую общность или абстрактность суждений, связанных с общесоциальной оценкой действительности, без значимой конкретизации данных оценок, что, собственно, сохраняет ценность исходного целевого запроса к нейросети.

3. LLM как интерпретатор научных фактов и тематических исследований в рамках информационных баз данных

М. В. Захарова отмечает, что обещаемые создателями ИИ-инструментов преимущества цифровизации исследований посредством внедрения ИИ в научный процесс концентрируются вокруг гипотетического избавления исследователей от рутинных задач и экономии времени, которое предполагается перенаправить на более существенные аспекты исследований, например, анализ, интерпретацию, обобщение данных (Захарова, 2024). В результате интерпретация становится одним из потенциальных векторов развития ИИ, а значит — и больших языковых моделей.

Как отмечает А. А. Бочкова: «Великие ученые — это не просто выдающиеся оптимизаторы обычных задач. Чаще всего великие ученые — это обычные оптимизаторы, решающие необычные оптимизационные задачи... Нормальная наука — это не просто оптимизация и не каталог оптимизационных задач, ожидающих решения. Фундаментальный барьер здесь тот же, что и для ученых в области ИИ: это концептуальная проблема постановки оптимизационной задачи» (Бочкова, 2025). Из концептуальной проблемы такой задачи проистекает актуальность той или иной научной идеи.

В рамках недавно проведенного исследования с задействованием «идейных агентов» на основе LLM и группы научных экспертов, целью которых была разработка инновационно-концептуальных положений области NLP, было выявлено, что идеи, сгенерированные ИИ, получили статистически значимо более высокие оценки по критериям новизны и увлекательности по сравнению с идеями людей. При этом по критерию осуществимости идеи ИИ оказались немного ниже, а по эффективности — немного выше, чем идеи людей, хотя эти различия не были статистически значимыми (Si et al., 2024).

Одновременно с положительными проявлениями LLM как средства научного познания и развития научного знания как такового выделим ряд проблем,

¹ Нейросеть оживила политиков. Дискуссия Сталина, Горбачева, Ельцина и Ленина на тему современная Россия. VC.RU. — URL: <https://vc.ru/id127146/780351-neiroset-ozhivila-politikov-diskussiya-stalina-gorbacheva-elcina-i-lenina-na-temu-sovremennaya-rossiya> (дата обращения: 26.07.2025).

которые препятствует представлению LLM как интерпретатора и концептуальной основы развития области научных открытий:

1. LLM могут с трудом генерировать точный контент для редких или малоизученных явлений, что создает проблемы для информирования об аномалиях или новых открытиях (Baryshnikov, 2024).

2. Некорректные и фальсифицированные библиографические ссылки, безусловно, пока ограничивают положительные стороны таких инструментов, как ChatGPT (Гребенщикова, 2024). Эту проблему подтвердил анализ 115 ссылок из 30 коротких статей по проблемам медицины, сгенерированных ChatGPT-3.5. Каждая статья содержала в среднем $3,8 \pm 1,1$ библиографических ссылок, которые были проверены в нескольких базах данных (Medline, Google Scholar и Directory of Open Access Journals). Среди этих ссылок 47% были сфабрикованы ИИ, 46% были подлинными, но неточными, и только 7% были правильными и точными (Bhattacharyya et al., 2023).

3. Эмпирический анализ, представленный в исследовании Я. Кузьмина и Е. Кручинской, показал, что LLM ошибаются не только в сложных заданиях, но и в элементарных. Подражая поведению Остапа Бендера, они подстраиваются под контекст, но не выявляют научной истины и не используют ее в ответах в качестве норматива (Кузьминов, Кручинская, 2024). Отсюда вытекает ключевой сдерживающий фактор для профессиональной продуктивизации инструментов ИИ.

4. Сетература как конструкт, использующийся в разработках М. А. Черняка и С. А. Морозовой в виде синонима к сетевой литературе, развивается на наших глазах уже в течение 20 лет и активно провоцирует филологов, философов, культурологов, специалистов цифровой гуманитаристики (Digital Humanities) вырабатывать особый терминологический аппарат (Черняк, Морозова, 2024). «Все основные понятия филологии: “читать”, “писать”, “текст”, “интерпретация” — приобретают новый смысл в электронной вселенной и подчас даже требуют другой терминологической артикуляции», — отмечает М. Эпштейн (Эпштейн, 2019, С. 16). Однако такая артикуляция должна находить надлежащее фактологическое подкрепление, а не становиться чем-то лингвистически сторонним. К примеру, автор поста из сообщества «Бюро переводов и школа переводчиков “Альянс ПРО”» социальной сети «ВКонтакте» отмечает следующее: «Еще одна болячка у LLM-ок — они любят изобретать слова, которые теоретически могли бы появиться у носителей, но тем не менее пока не появились»¹. В данном случае примечательным является справедливое замечание Б. В. Орехова: «По всей видимости, логическая процедура, выясняющая отчуждаемость знания от высказывания на формальном языке, состоит в установлении возможности создания

¹ Бюро переводов и школа переводчиков «Альянс ПРО»: Запись на стене vk-сообщества. — URL: https://vk.com/wall-72004531_8549 (дата обращения: 29.07.2025).

нарратива, который бы переформулировал такое высказывание, и в том, насколько такой нарратив равнозначен этому высказыванию. Если парафрастический нарратив возможен и если он равнозначен высказыванию на формальном языке, то и знание в естественных науках принципиально отчуждаемо от текста» (Орехов, 2024).

С целью поступательной разработки указанных проблемных положений в настоящее время создаются инструменты повышения качества обработки фактологических и интерпретируемых информационных материалов в архитектурах больших языковых моделей:

1. Разработана модель с говорящим названием «SAFE» (безопасность, Search-Augmented Factuality Evaluator, оценщик фактов с расширенным поиском). Используется большая языковая модель, чтобы разбить ответ на набор отдельных фактов и затем оценить точность каждого, отправив его на проверку в поисковик Google. Ответ на промпт (запрос) мог быть релевантным, нерелевантным, и каждый релевантный ответ оценивался как «поддерживаемый» и «неподдерживаемый» фактами. Также оценка качества модели учитывает процент подтвержденных фактов в ответе (точность) и долю ответов с предпочтительной длиной ответа для пользователя (Wei et al., 2024). Система «SAFE» позволяет постепенно вырабатывать подходы и методологии, которые могут послужить составными компонентами больших языковых моделей, генерирующих научный контент; при этом фактологическое подкрепление может служить косвенным инструментом повышения этического соответствия сгенерированных данных. Вместе с тем возможность работы с системами по типу «SAFE» способна служить не только средством фактологической поддержки систем с генерацией научных данных, но и средством популяризации повышения качества работы с этими данными через положительную или отрицательную фактологическую обусловленность — таким образом, пользователи моделей типа «SAFE» и возможных ее агентских исполнений могут со временем вырабатывать свой уникальный стиль работы с данными, имеющими фактологическую платформу, а значит, такой стиль может быть потенциально перенесен на системы генерации именно научного контента, и такой контент в перспективе может иметь более выраженную уникальность.

2. Исследователи из Университета Монаша создали искусственный интеллект, способный самостоятельно совершать научные открытия. Их программный продукт LLM4SD (Large Language Model for Scientific Discovery) обладает возможностью не только выполнения уже традиционных аналитических манипуляций с входящими информационными данными, но и способностью формирования в качестве надстройки над ними тех или иных гипотез. В отличие от существующих инструментов проверки гипотез, LLM4SD детально объясняет логику своих рассуждений. Прозрачность механизма

принятия решений позволяет исследователям проверить правильность выводов и избежать «галлюцинаций» — ошибочных заключений, которыми часто грешат нейросети (Zheng et al., 2023). Значимая степень проработки проблемы галлюцинаций может стать одним из ключевых моментов повышения степени доверия не только к самому программному продукту LLM4SD, но и ко всем его агентским и субагентским надстройкам — таким образом, рассмотренный программный продукт обладает потенциалом разработки перспективных научных направлений, построения самостоятельных систем и научных концепций как в прямом формате — методами исключительной программной генерации, так и в непрямом — когда на определенном этапе генерации вводятся пользовательские правки.

3. Последние достижения в области больших языковых моделей (LLM) значительно ускорили автоматизацию научных исследований, но в то же время вызвали серьезные опасения с точки зрения этики и безопасности. Чтобы систематически решать эти проблемы, была представлена модель SafeScientist — инновационная система для ученых, работающих с ИИ, которая специально разработана для повышения безопасности и этической ответственности при проведении научных исследований с использованием ИИ. SafeScientist заранее отклоняет задачи, которые не соответствуют этическим нормам или сопряжены с высоким риском, и уделяет особое внимание безопасности на протяжении всего исследовательского процесса (Zhu et al., 2025). Вопрос построения внешних систем оценки степени этичности научных данных является достаточно актуальным, поскольку, с одной стороны, системы генерации научной информации, наподобие LLM4SD, должны располагать внутренними инструментами оценки входных или сгенерированных научных данных на предмет их этической объективности; с другой стороны, внутренние алгоритмы этической оценки могут быть объектами воздействия генеративных манипуляционных техник, вследствие которых вместо генерации данных с этической обусловленностью отдается предпочтение генерации информации, наиболее соответствующей входящему пользовательскому запросу, которая не всегда может сочетать в себе высокую этическую оценку — в рамках этого свойства создание программных помощников формата SafeScientist объективна как для сохранения внутренней научной этики в границах функционирования систем генерации научного контента, так и для сохранения этики в ее этносоциальном содержании по отношению к различным социальным группам пользователей таких систем.

4. В то время как системы искусственного интеллекта демонстрируют экспоненциально растущие возможности, темпы самих исследований ИИ по-прежнему линейно ограничены когнитивными способностями человека, что создает все более серьезные препятствия для разработки. Была представлена модель ASI-ARCH как первая демонстрация искусственного

сверхинтеллекта для исследований ИИ (ASI4AI) в критической области исследования нейронной архитектуры — полностью автономная система, которая разрушает это фундаментальное ограничение, позволяя ИИ внедрять свои собственные архитектурные инновации и посредством этого формировать свои уникальные гипотезы (Liu et al., 2025). В результате модель построения архитектур, приближенных к сверхинтеллекту, с одной стороны становится альтернативным средством оценки эффективности когнитивных человеческих способностей, а с другой стороны выступает как собирательная архитектура больших языковых моделей, в рамках которых системы «сверхинтеллект» и «большая языковая модель» получают более выраженную сопряженность и взаимозависимость. Если рассматривать сверхинтеллект как альтернативу когнитивной модели человека, то в таком случае все созависимые термины со сверхинтеллектом, в частности AGI (Artificial General Intelligence) и ASI (Artificial Superintelligence), способны становиться средствами локального воплощения больших языковых моделей — по итогу, трансформация больших языковых моделей в AGI и ASI может быть взаимобратной в соответствии с контекстом сравнения этих систем.

Заключение

В соответствии с вышесказанным можно сделать вывод, что факт и интерпретация постепенно становятся средствами оценки информационных данных больших языковых моделей при формировании выходных результатов их работы, перемещаясь от трактовки личностного мнения к разработке научных данных из различных научно-информационных баз. Большие языковые модели расширяют возможности использования фактологических и информационных инструментариев с целью формирования потенциально нового научного дискурса. Кроме того, важность рассмотрения факта и интерпретации в системах с генерируемым контентом сопрягается не только с непосредственным получением новых информационных данных — в частности, на примере функционирования системы LLM4SD, — но также и данных, имеющих этическую обусловленность, как это было рассмотрено на примере работы SafeScientist. При этом этическая оценка может создаваться не только в условиях сопряженности с теми или иными категориями, но и как спектр информационных данных, выстроенных в рамках фактологической подкрепленности, — то есть на примере функционирования модели «SAFE». Этическая оценка работы больших языковых моделей может быть двухфакторной — с одной стороны, ее базис может строиться на прямой фактологической оценке генерируемых информационных или научных данных, а с другой стороны, надстройкой над таким базисом может служить оценка в прямых этических категориях с нужными локальными уточнениями, в зависимости от принадлежности сгенерированных данных к той или иной научной системе. Вопросы возможностей

обработки и разработки научных данных в рамках генерации соответствующего контента также сопрягаются с особенностями взаимосвязи систем «большая языковая модель» и «сверхинтеллект» (включая все связанные с ним системы, в особенности AGI и ASI), где структуры сверхинтеллекта являются как альтернативным подходом к оценке эффективности человеческого мышления, так и показателями эффективности локального использования агентов или архитектурных подвидов больших языковых моделей.

Список источников

Бочкова А. А. Искусственный интеллект: стратегии и методы решения сложных проблем // Надежность. — 2025. — Т. 25, № 1. — С. 46–57. — DOI 10.21683/1729-2646-2025-25-1-46-57. — EDN UGMEEC.

Гребенщикова Е. Г. Научные публикации в эпоху искусственного интеллекта // Научно-техническая информация. Серия 1: Организация и методика информационной работы. — 2024. — № 11. — С. 39–43. — DOI 10.36535/0548-0019-2024-11-5. — EDN LSYUTU.

Зайцев Д. В. Почему большие языковые модели не (всегда) рассуждают как люди? / Д. В. Зайцев // Вестник Московского университета. Серия 7: Философия. — 2024. — Т. 48, № 1. — С. 76–93. — DOI 10.55959/MSU0201-7385-7-2024-1-76-93. — EDN KEJEJL.

Захарова М. В. Интеллектуальные помощники для научного исследования в университетах / М. В. Захарова // Мир науки. Педагогика и психология. — 2024. — Т. 12, № 4. — EDN GGIFIX.

Кузьминов Я., Кручинская Е. Потенциал генеративного искусственного интеллекта для решения профессиональных задач // Форсайт. — 2024. — Т. 18, № 4. — С. 67–76. — DOI 10.17323/2500-2597.2024.4.67.76. — EDN CZHXZV.

Кузьминов В. Г., Швецов А. А. Искусственный интеллект: техно-лингвистический феномен или нечто большее? // Инновационные технологии в образовательной деятельности: Материалы XXVI Международной научно-методической конференции, Нижний

References

Bochkova A. A. Artificial intelligence: strategies and methods for solving complex problems. *Nadezhnost = Reliability*. 2025; 25(1): 46–57. <https://doi.org/10.21683/1729-2646-2025-25-1-46-57>. (In Russ.)

Grebenshchikova E. G. Scientific publications in the era of artificial intelligence. *Nauchno-tehnicheskaya informatsiya. Seriya 1: Organizatsiya i metodika informatsionnoy raboty = Scientific and technical information. Series 1: Organization and methodology of information work*. 2024; 11: 39–43. <https://doi.org/10.36535/0548-0019-2024-11-5>. (In Russ.)

Zaitsev D. V. Why do large language models not (always) reason like humans? *Vestnik Moskovskogo universiteta. Seriya 7: Filosofiya = Moscow University Bulletin. Series 7: Philosophy*. 2024; 48(1): 76–93. <https://doi.org/10.55959/MSU0201-7385-7-2024-1-76-93>. (In Russ.)

Zakharova M. V. Intelligent assistants for scientific research in universities. *Mir nauki. Pedagogika i psikhologiya = World of Science. Pedagogy and psychology*. 2024; 12(4). (In Russ.)

Kuzminov Ya., Kruchinskaya E. The potential of generative artificial intelligence for solving professional tasks. *Forsayt = Foresight*. 2024; 18(4): 67–76. <https://doi.org/10.17323/2500-2597.2024.4.67.76>. (In Russ.)

Kuzminov V. G., Shvetsov A. A. Artificial intelligence: techno-linguistic phenomenon or something more? *Innovatsionnyye tekhnologii v obrazovatel'noy deyatel'nosti: Materialy XXVI Mezhdunarodnoy nauchno-metodicheskoy konferentsii, Nizhniy Novgorod, 07 fevralya 2024 goda = Innovative technologies*

Новгород, 07 февраля 2024 года. — Нижний Новгород: Нижегородский государственный технический университет им. Р. Е. Алексеева, 2024. — С. 451–460. — EDN BPPVYO.

Орехов Б. В. Текст и знание в аспекте больших языковых моделей / Б. В. Орехов // Историческая информатика. — 2023. — № 4(46). — С. 104–113. — DOI 10.7256/2585-7797.2023.4.44180. — EDN BJBQBQ.

Сейновски Т. Антология машинного обучения. Важнейшие исследования в области ИИ за последние 60 лет. — М.: Бомбора, 2022. — 304 с. — ISBN 978-5-04-101347-9.

Соколыч В. Н. Открытая наука и современные научные публикации: этические требования и новые этические проблемы / В. Н. Соколыч // Труды БГТУ. Серия 6: История, философия. — 2024. — № 1(281). — С. 142–148. — DOI 10.52065/2520-6885-2024-281-27. — EDN BAENFW.

Черняк М. А., Морозова С. А. «Свет мой, GPT, скажи...», или Феномен художественного текста постлитературной эпохи / М. А. Черняк, С. А. Морозова // Мир русского слова. — 2024. — № 4. — С. 50–62. — DOI 10.21638/spbu30.2024.406. — EDN WRAQHM.

Эпштейн М. Будущее гуманитарных наук: Техногуманизм, креаторика, эротология, электронная филология и другие науки XXI века / Михаил Эпштейн. — М.: РИПОЛ классик / Панглосс, 2019. — 239 с. — (Серия «ЛекцииPRO»). ISBN 978-5-386-12499-1.

Baryshnikov P. What is scientific knowledge produced by Large Language Models? / P. Baryshnikov // Philosophical Problems of IT & Cyberspace (PhilIT & C). — 2024. — P. 89–103. 10.17726/philIT.2024.1.6.

Bhattacharyya M. High Rates of Fabricated and Inaccurate References in ChatGPT-Generated Medical Content / M. Bhattacharyya, V. M. Miller, D. Bhattacharyya, L. E. Miller // Cureus. 2023. — No. 15(5). — P. e39238. — DOI: 10.7759/cureus.39238.

Chalmers D. J. Does thinking require sensory grounding? From the history of philosophy to artificial intelligence / D. J. Chalmers // PhilArchive. 2023. — P. 22–45.

in educational activity: Proceedings of the XXVI international scientific and methodological conference. Nizhny Novgorod; 2024: 451–460. (In Russ.)

Orekhov B. V. Text and knowledge in the context of large language models. *Istoricheskaya informatika = Historical informatics*. 2023; 4(46): 104–113. <https://doi.org/10.7256/2585-7797.2023.4.44180>. (In Russ.)

Sejnowski T. The deep learning revolution: the most important AI research over the past 60 years. *M.: Bombora = Moscow: Bombora*, 2022; 304 p. ISBN 978-5-04-101347-9. (In Russ.)

Sokolchik V. N. Open science and modern scientific publications: ethical requirements and new ethical problems. *Trudy BGTU. Seriya 6: Istoriya, filosofiya = Proceedings of BSTU. Series 6: History, philosophy*. 2024; 1(281): 142–148. <https://doi.org/10.52065/2520-6885-2024-281-27>. (In Russ.)

Chernyak M. A., Morozova S. A. “Mirror, GPT, tell me...” or the phenomenon of literary text in the post-literary era. *Mir russkogo slova = World of Russian language*. 2024; 4: 50–62. <https://doi.org/10.21638/spbu30.2024.406>. (In Russ.)

Epstein M. The future of the humanities: techno-humanism, creatorics, erotology, electronic philology and other sciences of the XXI century. *M.: RIPOL klassik / Pangloss = Moscow: RIPOL Classic, Pangloss*, 2019; 239 p. ISBN 978-5-386-12499-1. (In Russ.)

Baryshnikov P. What is scientific knowledge produced by large language models? *Philosophical Problems of IT & Cyberspace*. 2024: 89–103. <https://doi.org/10.17726/philIT.2024.1.6>.

Bhattacharyya M., Miller V. M., Bhattacharyya D., Miller L. E. High rates of fabricated and inaccurate references in ChatGPT-generated medical content. *Cureus*. 2023; 15(5): e39238. <https://doi.org/10.7759/cureus.39238>.

Chalmers D. J. Does thinking require sensory grounding? From the history of philosophy to artificial intelligence. *PhilArchive*. 2023: 22–45.

Liu Y., Nan Y., Xu W., Hu X., Ye L., Qin Zh., Liu P. AlphaGo Moment for Model Architecture Discovery. 2025. arXiv. — URL: <https://arxiv.org/abs/2507.18074> (дата обращения: 24.07.2025).

Si Ch., Yang D., Hashimoto T. Can LLMs Generate Novel Research Ideas? A Large-Scale Human Study with 100+ NLP Researchers. 2024. arXiv. — URL: <https://arxiv.org/abs/2409.04109> (дата обращения: 24.07.2025).

Zheng Y., Koh H. Y., Ju J., Nguyen A. T. N., May L. T., Webb G. I., Pan Sh. 2023. arXiv. — URL: https://arxiv.org/abs/2310.07984?utm_source=Securitylab.ru (дата обращения: 24.07.2025).

Zhu K., Zhang J., Qi Z., Shang N., Liu Z., Han P., Su Y., Yu H., You J. SafeScientist: Toward Risk-Aware Scientific Discoveries by LLM Agents. 2025. arXiv — URL: <https://arxiv.org/abs/2505.23559> (дата обращения: 22.07.2025).

Liu Y., Nan Y., Xu W., Hu X., Ye L., Qin Zh., Liu P. AlphaGo Moment for Model Architecture Discovery. 2025. arXiv. — URL: <https://arxiv.org/abs/2507.18074>.

Si Ch., Yang D., Hashimoto T. Can LLMs Generate Novel Research Ideas? A Large-Scale Human Study with 100+ NLP Researchers. 2024. arXiv. — URL: <https://arxiv.org/abs/2409.04109>.

Zheng Y., Koh H. Y., Ju J., Nguyen A. T. N., May L. T., Webb G. I., Pan Sh. P. 2023. arXiv. — URL: https://arxiv.org/abs/2310.07984?utm_source=Securitylab.ru.

Zhu K., Zhang J., Qi Z., Shang N., Liu Z., Han P., Su Y., Yu H., You J. SafeScientist: Toward Risk-Aware Scientific Discoveries by LLM Agents. 2025. arXiv — URL: <https://arxiv.org/abs/2505.23559>.

Для цитирования: Филимонов В. Ю. Научные открытия методами LLM: факты и интерпретация. // Гуманитарий Юга России. — 2026. — Т. 15. — № 2 (78). — С. 161–172. DOI 10.18522/2227-8656.2026.2.12 EDN KQERXA

История статьи:
Поступила в редакцию — 13.12.2025
Одобрена после
рецензирования — 05.03.2026
Принята к публикации — 13.03.2026

Сведения об авторе

Филимонов Владимир Юрьевич
Аспирант 3 курса Пятигорского
государственного университета
SPIN-код: 2080-4910
AuthorID РИНЦ: 1199725
fil2016vo@yandex.ru

Information about author

Vladimir Yu. Filimonov
3rd year postgraduate student,
Pyatigorsk State University
WoS ResearcherID: PBV-2236-2025
fil2016vo@yandex.ru