

Историческая информатика

Правильная ссылка на статью:

Орехов Б.В. Текст и знание в аспекте больших языковых моделей // Историческая информатика. 2023. № 4. С. 104-113. DOI: 10.7256/2585-7797.2023.4.44180 EDN: BJQBQB URL: https://nbpublish.com/library_read_article.php?id=44180

Текст и знание в аспекте больших языковых моделей

Орехов Борис Валерьевич

ORCID: 0000-0002-9099-0436

кандидат филологических наук

старший научный сотрудник, Лаборатория цифровых исследований литературы и фольклора,
Институт русской литературы (Пушкинский Дом) РАН

119331, Россия, Москва, г. Москва, ул. Крупской, 13, кв. 77

✉ nevmenandr@gmail.com



[Статья из рубрики "Искусственный интеллект и наука о данных"](#)

DOI:

10.7256/2585-7797.2023.4.44180

EDN:

BJQBQB

Дата направления статьи в редакцию:

30-09-2023

Дата публикации:

31-12-2023

Аннотация: В центре внимания этого текста – влияние больших языковых моделей на самоопределение гуманитарной науки. Большие языковые модели умеют порождать правдоподобные тексты. Кажется, что они, таким образом, становятся в один ряд с другими инструментами, которые на протяжении развития техники и технологий, избавляли человека от рутины. В то же время для гуманитарных наук очень велика индивидуализированность порождаемых текстов, и само знание оказывается тесным образом связано с его текстовым воплощением. В случае, если мы соглашаемся с тем, что знание – это текст, и воплощенное в другом тексте перед нами предстает уже иное знание, значит, гуманитарной науке придется ответить на вопрос, каким образом текст, порожденный человеком, ценностно отличается от такого же текста, порожденного машиной. В тексте работы ставятся методологические и эпистемологические проблемы

соотношения текстов естественного и искусственного происхождения в том случае, если они выполнены в жанре научного произведения. Разница между такого рода артефактами ясно видна только для некоторых научных дисциплин, а в отношении остальных вызывает вопросы. Эти вопросы должны быть разрешены с помощью глубокой рефлексии, которая не так остро требовалась в последние века развития гуманитарной науки, но которая потребуется теперь от ученого-гуманитария. Гуманитарию придется эксплицитно противопоставить себя большим языковым моделям и доказать значимость его работы по сравнению с тем, что может сгенерировать нейросеть.

Ключевые слова:

большие языковые модели, chatgpt, научные публикации, методология науки, текстовые генераторы, знание, наука, текст, формальные языки, гуманитарные науки

Введение

«Историки пишут тексты», — так начинает свою статью В. Канштайнер [\[12\]](#). И в этом утверждении действительно ключевая проблема взаимоотношений искусственного интеллекта (ИИ) и историка, и даже шире — любого ученого из стана гуманитариев.

Единственное затруднение, которое видится В. Канштайнеру в современных, еще не вполне совершенных больших языковых моделях (LLM), состоит в том, что они не умеют отличать истину от лжи, но и это затруднение должно иметь техническое решение в перспективе соединения текстового генератора с базами знаний. Если принять этот тезис, то основной нерв дискуссии будет лежать где-то в других сферах, например, в этичности LLM. Поэтому автор статьи довольно быстро переключается на моральные вопросы, строя мысленные эксперименты, в которых участвует Гитлер, память о Холокосте и подобные им эмоционально нагруженные проблемы.

Кстати, когда Канштайнер касается проблемы правды, он говорит о ней так, как будто само по себе это понятие не проблематично. Собственно, в истории, как показывает Ф. Анкерсмит [\[1\]](#), возможно бесконечное число нарративов, и ни один из них не может обоснованно претендовать на роль «правды». Это тоже давно известная и получившая свою рефлексия проблема.

В то же время, как кажется, появление LLM ставит перед ученым-гуманитарием более сложные концептуальные вопросы самоопределения, которые раньше в этом поле не звучали. Они затрагивают самые базовые основания научной деятельности, и именно успехи нейросетей провоцируют такие вопросы задавать.

Действительно, «историки пишут тексты». Тексты пишут и представители других наук. Но стереотипные образы ученого и науки выводят на первый план совсем иное: не текст, а знание. Именно знание и истина становятся наиболее значимыми результатами научной деятельности для взгляда со стороны.

Сама система ChatGPT так определяет науку: «Наука — это систематический и организованный подход к изучению природы, общества и различных явлений, основанный на наблюдении, экспериментах и логическом анализе данных. Цель науки — расширение знаний и понимания мира, создание новых технологий и решение практических проблем». Здесь не упоминаются тексты. Если следовать этому определению, тексты для ученых как будто не являются главным продуктом их работы.

Главное в науке как деятельности — это получение нового знания. В предельном смысле — научные открытия.

Вопрос отчуждаемости знания от текста

В идеализированном схематичном мире ученый действительно производит знание, которое не зависит от текстового воплощения. Г. Гачев цитирует К. Гаусса: «Пока мы продумываем предмет непосредственно только в своем представлении, мы не нуждаемся ни в именах, ни в знаках. Они становятся необходимыми только тогда, когда мы хотим объяснить его другим» [\[3\]](#).

В реальности отношения между ученым и его продуктом сложнее. Сама оценка работы ученого — это его публикационная активность. Поскольку невозможно формально и строго оценить значимость научной идеи (и даже ее наличия), все, что могут измерять администраторы, это число и статус обнародованных текстов. Принцип «Publish or perish» предельно конкретно иллюстрирует важность факторов, влияющих на трудоустройство ученого и успешность его научной карьеры.

В таком случае главный вопрос, который ставит перед ученым появление LLM, это вопрос о том, способен ли ученый делать что-то такое, в чем LLM не могут его повторить.

Ученые могут порождать тексты, LLM тоже умеют порождать тексты. Но создает ли ученый то самое знание, о котором идет речь в определениях науки? И далее: существует ли знание, поиском которого должна заниматься наука? Если текст, в котором воплощается знание, и само знание неразделимы, то LLM вполне могут повторить и, следовательно, заменить ученых, производя тексты. Возможно, знание и истина — это мифы, и ученые всего лишь имитируют их поиск, и в этом смысле ничем не отличаются от LLM, основной деятельностью которых является имитация? Неприятный характер этих вопросов в том, что они имеют институциональные следствия: стоит ли выделять бюджеты на имитационную деятельность людей, если с ней успешно справляются более дешевые языковые модели?

В нарисованной схеме представители гуманитарных и естественных наук как будто находятся в разном положении. Имитировать гуманитарный текст проще, чем естественнонаучный. Однако и для естественных наук ответы на названные вопросы не вполне однозначны. Благодаря работам Б. Латурса [\[7\]](#) и из-за осознания кризиса воспроизводимости в естественных науках мы знаем, что исследование трудно отделить от лаборатории, в которой оно было выполнено, а, следовательно, и от текста, который создали сотрудники этой лаборатории для последующей публикации. Отчуждаемо ли то, что сказано в научной статье по химии или физике, от текста этой научной статьи?

Любой ответ на этот вопрос будет дискуссионным, но симптоматично стремление рассматривать такие научные произведения как текстоцентричные, иными словами, рассматривать их таким способом, при котором особенности текстового воплощения являются частью постулируемого знания.

В книге Г. Гачева об этом говорится довольно определенно, и наука в аспекте формы сопоставляется с художественной литературой: «наука берется <...> не только как сумма идей, опытов, знаний, но и как научная литература, а каждое научное сочинение рассматривается не только как изложение определенных взглядов на явление природы, но и как текст на естественном языке, где образ и стиль имеют фундаментально-определяющее значение для воссоздаваемой картины мира и построений теории» [\[3\]](#). В схожем русле рассуждает о риторике научных статей Латурс, показывая, к каким

текстовым приемам прибегают их авторы, стремясь к максимальному воздействию на читателя [\[7\]](#). О том, как тесно текст соотносится с наукой, говорят и менее титулованные, но не менее вдумчивые авторы: «Письмо играет такую центральную роль в познании, обучении и исследованиях, что удивительно, как мало мы о нем думаем» [\[6, с. 9\]](#).

Сейчас, благодаря LLM, мы вправе задать себе вопрос, не является ли наука не просто деятельностью, на которую оказывают (дополнительное) влияние образ и стиль, но тем, что принципиально не отличается от художественной литературы с ее образами и стилем.

Естественные науки пользуются формальными языками, например, статья по химии может содержать формулы химической реакции. Но LLM также могут выучиться формальным языкам и порождать на них высказывания, как это уже произошло с языками естественными. Химические формулы можно генерировать в ответах моделей уже сейчас. По всей видимости, логическая процедура, выясняющая отчуждаемость знания от высказывания на формальном языке, состоит установлении возможности создания 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. Как мы уже говорили, и здесь существуют сложности, описанные Б. Латуром, но они будут иметь характер частных соображений.

Если нарратив не будет соответствовать высказыванию на формальном языке, значит, проблема статуса знания не касается сугубо гуманитариев, а общая для всех наук. Иными словами, представители естествонаучного сообщества точно так же зависимы в результатах своей деятельности от текста, как и представители гуманитарного сообщества.

Есть в разговоре на эту тему и плоскость, в которой исторические науки оказываются ближе к естественным, чем к гуманитарным, — это соотнесенность с реальностью. Если в статье на химическую или физическую тематику будет упоминаться эксперимент, которого на самом деле никто не проводил, такой казус будет примерно соответствовать упоминанию в историческом исследовании «факта», не имеющего подтверждения в источниках. Такого рода случаи подробно разбирает в своем тексте В. Канштайнер, но для нас важно, что это обстоятельство существенно для демаркации наук разного типа: в литературоведческом анализе обнаружить такие «факты», имеющие бинарный статус («истинно» или «ложно»), гораздо труднее. Трактовки художественных образов, их встраиваемость в системы более высоких порядков (образные системы и мотивные структуры циклов, авторских корпусов, корпусов эпохи и стиля) могут быть только более или менее убедительными. Убеждение — категория риторики, а не знания, а риторика актуализируется только в конкретном тексте.

Н. Луман так описывает историю подходов к новому знанию: «До Куна все более ранние описания мира, которые не соотносились с последними исследованиями, рассматривались как более-менее провальные попытки получить научное знание» [\[13, с. 10-11\]](#). Благодаря Куну «устаревшие» теории стали рассматриваться как происходящие из других парадигм, а, значит, не как неправильные, а как альтернативные. «Можно только сказать: мы имеем дело с иной парадигмой, чьи претензии на первенство в могут быть сформулированы только ее собственными терминах» [\[13, с. 11\]](#). По большому счету, особенность формулировки — это тоже категория риторики.

Невозможно рассматривать как истинное или ложное высказывание вроде такого: «суть поэтического мира Пастернака — это быт и экстатичность» [11, с. 210] или такого: «И<нфинитивное >П<исьмо> трактует о некой виртуальной реальности, которую поэт держит перед мысленным взором, о некоем “там”, в отличие от другого минималистского стиля—назывного, который рисует описываемое как имеющее место здесь и сейчас» [5, с. 250].

На этом примере мы видим, что знание, которое производят ядерные гуманитарные науки (то есть в данном случае не рассматриваем пограничные случаи вроде лингвистики или истории), неотчуждаемо от текста. Глубоко индивидуализированные работы аналитического характера выдают источник аргументов каждого из процитированных ученых. Для К. Тарановского это классическая герменевтическая традиция, а для А. К. Жолковского — структурная лингвистика с ее опорой на функциональность частей речи.

Забытые литературоведческие работы советского времени также подтверждают тесную связь текста с гуманитарным знанием. Эти труды оказались забыты именно по той причине, что они не соотносимы с актуальной идеологической повесткой (в некотором роде аналогом куновской «парадигмы» в гуманитарной науке). Ср. текст аннотации современной книги о русском реализме: «Научная дискуссия о русском реализме, скомпрометированная советским литературоведением, прервалась в постсоветскую эпоху» [8].

Даже позитивистское знание биографического или текстологического характера не существует отдельно от своей аргументации, а значит, существует внутри того же самого текста. Так, аргументация датировки стихотворения Тютчева «Давно ль, давно ль, о Юг блаженный...» связывает факты непосредственных наблюдений поэтом, отрицая его право на фантазию: «К. В. Пигарев предположил, что это стихотворение “написано в декабре 1837 г., по возвращении из Генуи в северную Италию, в Турин”. Согласиться с этим предположением невозможно. *Точность пейзажных зарисовок в лирике Тютчева хорошо известна, и с этим нельзя не считаться при датировке его стихов.* В данном же случае предположению К. В. Пигарева противоречит весь текст стихотворения, где “блаженному” Югу, “лазоревой равнине” его моря противопоставлен суровый Север <...>. Этот пейзаж ни в коей мере не соответствует пейзажу Турина и его окрестностей, расположенных в предгорьях Альп, на одной широте с Генуей (хотя, разумеется, климат этих мест в сравнении с мягким климатом Генуи кажется более суровым)» [4, с. 288]. Здесь курсивное выделение мое — Б. О. Речь идет о фальсифицируемом историческом знании, но способ его установления опирается на неконвенциональную категорию убедительности.

Но все же в отличие от естественных наук, где основанием для текста является эксперимент, в гуманитарных науках такая эмпирическая база отсутствует, и поэтому отделимость научного текста от знания по оси фактической достоверности представляется еще более проблематичной. До современной ситуации, включающей фактор нейросетей, эта проблема не ощущалась настолько остро.

Показательна в этом смысле рефлексия М. Л. Гаспарова над трактовками изгнания Овидия из Рима: «Гипотез о том, в чем состоял “проступок” Овидия, за пять с лишним веков филологической науки накопилось столько, что недавний обзор их занял довольно толстую книгу, а приложенный к ней далеко не полный перечень насчитывает 111 аргументированных мнений. При этом в разные эпохи любопытнейшим образом

преобладали варианты разных гипотез. Первый период — средние века и Возрождение: комментаторы Овидия еще не располагают никаким материалом, кроме овидиевских текстов и собственной фантазии, а фантазия эта небогата. Всему виной — языческое распутство: если Овидий что-то совершил, то это было прелюбодеяние с женой или дочерью императора, если Овидий что-то увидел, то это был император, предающийся содомии или кровосмешению с родной дочерью, а может быть, императрица, выходящая из купальни. Второй период — к XVIII в. историки разбираются в лицах и датах, открывается одновременность ссылки Овидия и Юлии, является новая версия. Всему виной — любовная история: если Овидий виновен делом, то он был любовником Юлии Младшей, если виновен взглядом, то он был свидетелем, а может быть, и пособником любви Юлии и Силана. Третий период — трезвый XIX век смещает интерес с романического аспекта событий на политический: Овидий пострадал за то, что участвовал в заговоре (или хотя бы знал о нем), который будто бы организовали против Августа Юлия и Эмилий Павел с целью возвести на престол Агриппу Постума. Четвертый период — в конце XIX в. пробуждается внимание к темной, иррациональной стороне античного мира: неизреченная вина Овидия оказывается не политической, а религиозной, он то ли нарушил устав каких-то неразглашаемых таинств (в честь Исиды, в честь элевсинской Деметры, в честь римской Доброй Богини), то ли участвовал в магических гаданиях о судьбе императора. Наконец, наступает пятый период — и XX век, переживший фашизм и другие формы тоталитаризма, говорит: никакого проступка вообще не было названо, Овидию сказали: «ты виноват — тебя наказывают; а в чем виноват — ты сам должен понимать»; и все покаяния Овидия так невразумительны именно потому, что он сам не знает, в чем он виноват» [\[2\]](#). Здесь филолог обращает внимание на относительность гуманитарного знания, его зависимость от идеологического контекста, и конечную сводимость к той же категории убедительности: «Все эти теории (кроме, разве, самых ранних) почти правдоподобны, но ни одна из них не убедительна до конца».

Следует уточнить, что такой текстоцентричный статус гуманитарной науки не всегда был обязательным. Для Платона, как следует из «Федра», форма диалога была не окончательной формой выражения. Учение существовало в устной форме, а книга считалась уязвимой, поскольку не могла за себя постоять в философском споре, всегда говорила одно и то же (философ может варьировать тезисы), и не была ориентирована на собеседника (письменный текст обращается сразу ко всем, а философ — к конкретному адресату, и поэтому действует тоньше и острее) [\[10\]](#). Таким образом, знание было отчуждаемо от текста, и письменный текст не представляется лучшей формой его репрезентации.

В некотором смысле можно трактовать следующий пассаж в «Евтидеме» (289b и далее) как критику LLM:

— Следовательно, прекрасный мой мальчик, — продолжал я, — мы нуждаемся в таком знании, в котором сочеталось бы умение что-то делать и умение пользоваться сделанным.

— Это ясно, — отвечал он.

— Значит, как видно, нам вовсе не нужно становиться искусными в изготовлении лир и сноровистыми в подобном умении. Ведь здесь искусство изготовления и искусство применения существуют порознь, хотя и относятся к одному и тому же предмету, ибо искусство изготовления лир и искусство игры на них весьма отличаются друг от друга. Не так ли?

— Но, во имя богов, — сказал я, — если мы изучим искусство составления речей, то именно приобретение этого искусства сделает нас счастливыми?

— Я этого не думаю, — отвечал Клиний, схватив мою мысль.

— А как ты можешь это обосновать? — спросил я.

— Я знаю некоторых составителей речей, не умеющих пользоваться собственными речами, которые сами они сочинили, подобно тому как изготовители лир не умеют пользоваться лирами. В то же время есть другие люди, умеющие пользоваться тем, что первые приготовили, хотя сами приготовить речи не умеют. Ясно, что и в деле составления речей искусство изготовления — это одно, а искусство применения — другое.

Платон говорит об искусстве составления речей (ἡ Λογοποιικὴ τέχνη, 289c 7), которое фактически и составляет суть умений LLM. При этом невозможность достичь счастья — это и есть для философа основной способ критики.

Как критику следует воспринимать и иронию Дж. Свифта: «Но мир вскоре оценит всю полезность этого проекта; и он льстил себя уверенностью, что более возвышенная идея никогда еще не зарождалась ни в чьей голове. Каждому известно, как трудно изучать науки и искусства по общепринятой методе; между тем благодаря его изобретению самый невежественный человек с помощью умеренных затрат и небольших физических усилий может писать книги по философии, поэзии, политике, праву, математике и богословию при полном отсутствии эрудиции и таланта»^[9].

Заключение

Отчуждаемость знания от текста представляет собой принципиальную проблему, которую, почти коснувшись, все же не заметил В. Канштайнер. Сама по себе она не нова, и уже осознана гуманитариями (отчасти это видно по тексту М. Л. Гаспарова об изгнании Овидия). Но благодаря LLM ее статус меняется. Из абстрактно-теоретического он превращается в конкретно-методологический.

Нейросети прежде всего ставят перед гуманитарной наукой вопрос о том, что есть в ней помимо текста, есть ли у знаков в научном тексте референт. Возможно, что уже в ближайшем будущем наличие референта ученым-гуманитариям придется показывать и доказывать для внешнего наблюдателя.

Пока что LLM плохо анализируют тексты (например, поэтические): вместо ожидаемых ориентированных на конкретное произведение этюдов они генерируют последовательности общих слов, приложимых к произвольному сочинению. Возможно, что с увеличением числа параметров и с жанровой специализацией эта проблема будет преодолена, а превосходство человека в этой области продлится ненадолго.

Тем не менее, индивидуализированность гуманитарных текстов может быть важным механизмом, отличающим сгенерированный текст от естественного. Нейросети продуцируют «среднее арифметическое» от обучающих данных. То есть сама технология принципиально противоречит идеи оригинальности. Парадоксальным образом оригинальность проявляет себя тогда, когда модель допускает ошибки. Из этого следует, что чем лучше обучена LLM, тем менее она оригинальна.

Оригинальность и соотносимая с ней гениальность ученого — это мифология эпохи

Романтизма. Романтизм навязывает обществу миф о гениальном художнике, музыканте, поэте и учёном, которые преодолевают замшелость взглядов своих современников. Замешательство перед успехом нейросетей — это замешательство перед тем, что миф не оправдывается, перед тем, что электронное устройство, лишённое сознания, повторяет то, на что, согласно мифу, был способен только гениальный человек. Этим в том числе объясняется то, что успех LLM произвел значительный медийный эффект.

Возможно, LLM заставят гуманитариев вернуться к романтическому мифу индивидуализированной науки и опереться на него в борьбе с цифровым соперником. Ведь если способностей языковой модели достаточно для того, чтобы имитировать научный текст, то это означает, что наука — это всего лишь технология, «среднее арифметическое» от прочитанного. А каждый гуманитарий знает, что это не так, и находится в постоянном поиске приращения смысла.

Библиография

1. Анкерсмит Ф. Р. История и тропология. Взлет и падение метафоры. М.: Прогресс-Традиция, 2003.
2. Гаспаров М. Л. Овидий в изгнании // Публий Овидий Назон. Скорбные элегии. Письма с Понта. М.: Наука, 1978. С. 189-224.
3. Гачев Г. Д. Наука и национальная культура (гуманитарный комментарий к естествознанию). Ростов-на-Дону: Изд-во Рост. ун-та, 1993.
4. Динесман Т. Г. О датировках и адресатах некоторых стихотворений Тютчева // Летопись жизни и творчества Ф. И. Тютчева. [М.]: ООО "Литограф"; [Мураново]: Музей-усадьба "Мураново" им. Ф. И. Тютчева, 1999—2012. Кн. 1: 1803—1844. 1999. С. 277-290.
5. Жолковский А. К. Инфинитивное письмо: тропы и сюжеты (Материалы к теме) // Эткиндовские чтения: сборник статей по материалам Чтений памяти Е. Г. Эткинда (27—29 июня 2000 г.). СПб., 2003. С. 250-271.
6. Зонке А. Как делать полезные заметки. Эффективная система организации идей по методу Zettelkasten. М.: Манн, Иванов и Фербер, 2022.
7. Латур Б. Наука в действии. СПб.: Издательство Европейского университета в Санкт-Петербурге, 2013.
8. Русский реализм XIX века. Общество, знание, повествование. М.: Новое литературное обозрение, 2020.
9. Свифт Дж. Собрание сочинений: В 3 т. Т. 1: Сказка бочки; Путешествия Гулливера. М.: ТЕРРА — Книжный клуб, 2008.
10. Слезак Т. А. Как читать Платона. СПб.: Изд-во Санкт-Петербургского ун-та, 2008.
11. Тарановский К. О поэзии и поэтике. М.: Языки русской культуры, 2000.
12. Kansteiner W. Digital Doping For Historians: Can History, Memory, And Historical Theory Be Rendered Artificially Intelligent? // History and Theory. 2022. Vol. 61. No. 4. Pp. 119-133.
13. Luhmann N. The modernity of science // New German Critique. 1994. No. 61. Pp. 9-23.

Результаты процедуры рецензирования статьи

В связи с политикой двойного слепого рецензирования личность рецензента не раскрывается.

Со списком рецензентов издательства можно ознакомиться [здесь](#).

Последние несколько десятилетий оказались ознаменованы растущей междисциплинарностью в науке. В самом деле, чего стоит только биоэтика, разработанная В.Р. Поттером, который прозорливо заметил важность этического

осмысления новых медицинских технологий. Однако не меньшую роль применимо к историческим наукам играют современные информационно-коммуникативные технологии: здесь мы имеем в виду историческую информатику, первые работы в рамках которой в Советском Союзе начались еще в 1970-е гг.

Указанные обстоятельства определяют актуальность представленной на рецензирование статьи, предметом которой являются большие языковые модели. Автор ставит своими задачами проанализировать вопрос отчуждаемости знания от текста, рассмотреть проблему нейросетей в контексте гуманитарных наук, а также выявить причины замешательства современников перед нейросетями.

Работа основана на принципах анализа и синтеза, достоверности, объективности, методологической базой исследования выступает системный подход, в основе которого находится рассмотрение объекта как целостного комплекса взаимосвязанных элементов. Научная новизна статьи заключается в самой постановке темы: автор стремится охарактеризовать проблему взаимоотношений искусственного интеллекта и историка.

Рассматривая библиографический список статьи как позитивный момент следует отметить его разносторонность: всего список литературы включает в себя 13 различных источников и исследований. Несомненным достоинством рецензируемой статьи является привлечение зарубежной англоязычной литературы, что усиливает и научную новизну. Из используемых автором исследования укажем на труды В. Канштайнера и Н. Лумана, в центре внимания которых различные аспекты изучения подходов к новому знанию. Заметим, что библиография обладает важностью как с научной, так и с просветительской точки зрения: после прочтения текста статьи читатели могут обратиться к другим материалам по ее теме. В целом, на наш взгляд, комплексное использование различных источников и исследований способствовало решению стоящих перед автором задач.

Стиль написания статьи можно отнести к научному, вместе с тем доступному для понимания не только специалистам, но и широкой читательской аудитории, всем, кто интересуется как влиянием нейросетей на развитие гуманитарных наук, в целом, так и взаимосвязью текста и нового знания, в частности. Апелляция к оппонентам представлена на уровне собранной информации, полученной автором в ходе работы над темой статьи.

Структура работы отличается определенной логичностью и последовательностью, в ней можно выделить введение, основную часть, заключение. В начале автор определяет актуальность темы, показывает, что «главный вопрос, который ставит перед ученым появление LLM, это вопрос о том, способен ли ученый делать что-то такое, в чем LLM не могут его повторить». Автор обращает внимание на то, что «в отличие от естественных наук, где основанием для текста является эксперимент, в гуманитарных науках такая эмпирическая база отсутствует, и поэтому отделимость научного текста от знания по оси фактической достоверности», при этом до эпохи нейросетей данная проблема не виделась настолько актуальной. В работе показано, что благодаря LLM из абстрактно-теоретического проблема отчуждаемости знания от текста превращается в конкретно-методологическую. Справедливо авторское мнение о том, что «нейросети прежде всего ставят перед гуманитарной наукой вопрос о том, что есть в ней помимо текста, есть ли у знаков в научном тексте референт».

Главным выводом статьи является то, что «индивидуализированность гуманитарных текстов может быть важным механизмом, отличающим сгенерированный текст от естественного».

Представленная на рецензирование статья посвящена актуальной теме, вызовет читательский интерес, а ее материалы могут быть использованы как в учебных курсах, так и в рамках изучения проблем взаимодействия искусственного интеллекта и ученого-

гуманитария.

В целом, на наш взгляд, статья может быть рекомендована для публикации в журнале «Историческая информатика».