



DOI: 10.22363/2313-2299-2026-17-2-496-516

EDN: LKJCAT

Научная статья / Research Article

Произведения больших языковых моделей как симулякр смысла. Семантическая энтропия и особенности новейших алгоритмов обнаружения использования искусственного интеллекта

М.Л. Новикова , Ф.Н. Новиков 

Российский университет дружбы народов, Москва, Российская Федерация
[✉novikova-ml@rudn.ru](mailto:novikova-ml@rudn.ru)

Аннотация. Ситуация, сложившаяся в связи со стремительным развитием и повсеместной доступностью искусственного интеллекта в целом и больших языковых моделей в частности, способных генерировать качественные, связные, правдивые или правдоподобные тексты на различных языках, является актуальной проблемой. Путем исследования таких междисциплинарных феноменов на стыке, как «семантическая энтропия», «симулякр смысла» и «перплексия», опираясь на теоретические базы количественной лингвистики, философии и информационных технологий, авторы статьи выдвигают гипотезу о том, что следствием таких изменений является радикальная трансформация информационной картины мира. Данная ситуация, с одной стороны, ведет к глобальной семантической девальвации и скептическому отношению к вновь создаваемым произведениям письменной речи, а с другой стороны повышает ценность аутентичных текстов, созданных человеком (подкрепленных его авторитетом и фактической базой). В статье также рассматриваются особенности новейших технических и семантических механизмов обнаружения фактов использования нейросетей при создании текстов и делается попытка описать дальнейшие тенденции в этой области. Авторами рассматривается феномен генеративного искусственного интеллекта (LLM) через призму концепции симулякра Жана Бодрийяра. Анализируется техническая природа нейросетей как вероятностного распределения токенов, создающего иллюзию семантической связности. На примерах демонстрируется разрыв между синтаксической корректностью и отсутствием интенциональности (смыслополагания), что позволяет квалифицировать продукты ИИ как «чистую форму», лишенную референта в реальности.

Ключевые слова: искусственный интеллект, семиотика, семантика, симулякр смысла, перплексия

Вклад авторов: вклад авторов равнозначен на всех этапах исследования и подготовки текста статьи.

Конфликт интересов: авторы заявили об отсутствии конфликтов интересов.

История статьи: дата поступления: 21.11.2025; дата приема в печать: 5.12.2025.

© Новикова М.Л., Новиков Ф.Н., 2026



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License
<https://creativecommons.org/licenses/by-nc/4.0/legalcode>

Для цитирования: Новикова М.Л., Новиков Ф.Н. Произведения больших языковых моделей как симулякр смысла. Семантическая энтропия и особенности новейших алгоритмов обнаружения использования искусственного интеллекта // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2026. Т. 17. № 2. С. 496–516. <https://doi.org/10.22363/2313-2299-2026-17-3-496-516> EDN LKJCAT

Works of Large Language Models as Simulations of Meaning. Semantic Entropy and Features of the Latest AI Detection Algorithms

Marina L. Novikova  , Philipp N. Novikov 

RUDN University, Moscow, Russian Federation

novikova-ml@rudn.ru

Abstract. The situation created by the rapid development and widespread availability of artificial intelligence in general, and large language models in particular, which can generate high-quality and coherent true or truthful texts in different languages, is a pressing problem. By studying interdisciplinary phenomena at the junction such as “semantic entropy”, “simulation of meaning” and “perplexity”, based on the theoretical bases of quantitative linguistics, philosophy and information technology, the authors hypothesize that the consequence of such changes is a radical transformation of the informational picture of the world. This situation, on the one hand, leads to a global semantic devaluation and skepticism towards newly created written speech works, and, on the other hand, increases the value of authentic human-made texts (supported by their authority and factual basis). The article also discusses the features of the latest technical and semantic mechanisms for detecting the use of neural networks in text creation, and attempts to describe further trends in this area. The authors consider the phenomenon of generative artificial intelligence (LLM) through the prism of the concept of Jean Baudrillard’s simulation. The technical nature of the neural networks is analyzed as a probabilistic distribution of tokens, creating an illusion of semantic coherence. The examples show a gap between syntactic correctness and the lack of intentionality (meaning-making), which allows AI products to be classified as “pure form” that lacks a referent in reality.

Keywords: artificial intelligence, semiotics, semantics, simulation of meaning, perplexity

Authors’ contribution: The authors contributed equally to this article and preparation of the article text.

Conflict of interests: The authors stated that there were no conflicts of interest.

Article history: Received: 22.11.2025; accepted: 5.12.2025.

For citation: Novikova, M.L., & Novikov, Ph.N. (2026). Work of Large Language Models as Simulations of Meaning. Semantic Entropy and Features of the Latest AI Detection Algorithms. *RUDN Journal of Language Studies, Semiotics and Semantics*, 17(2), 496–516. (In Russ.). <https://doi.org/10.22363/2313-2299-2026-17-3-496-516> EDN LKJCAT

Введение.

Искусственный интеллект и лингвистика: точки сопряжения

Лингвистическая мысль в различные периоды развития языкознания отвечает на актуальные вызовы времени. Использование искусственного

интеллекта в лингвистике по праву считается и действительно является одной из самых активно развивающихся областей современной науки, что связано с доминированием статистических методов, нейросетевого моделирования и дистрибутивной семантики. Искусственный интеллект «существует как *динамическое поле возможностей (курсив наш)*, которое актуализируется по-разному в зависимости от контекста взаимодействия, от языка описания, от самой постановки вопроса» [1. С. 326].

На протяжении десятилетий лингвистика и информатика двигались навстречу друг другу — от формальных грамматик Н. Хомского, которые легли в основу языков программирования, до современных нейронных сетей, способных генерировать информацию в любой цифровой форме.

В середине XX в. два направления исследования переживали свое рождение практически одновременно: искусственный интеллект (в первоначальном понимании слова) и генеративная лингвистика. В центре этого интеллектуального взрыва, качественного изменения в исследованиях находился Ноам Хомский, а его книга «Синтаксические структуры» [2] — формализованная общая теория лингвистической структуры, база современных теорий, необходимых для решения задач, возникающих при автоматической переработке естественно-языковой информации, не просто обогатила лингвистику, но и предоставила математический аппарат, необходимый компьютерным наукам для работы с языком. Во второй половине XX в. доминирующим был символичный искусственный интеллект. Пытаясь реализовать эти идеи, ученые выдвинули гипотезу о том, что описание всех правил грамматики (следуя Н. Хомскому), создание огромных «деревьев» правил, загруженных в компьютер, даст возможность машине овладеть языком. Но язык — подвижная динамичная структура с нарушениями правил; жесткие грамматики Н. Хомского «ломались» на реальных текстах. Грамматика Н. Хомского не только описывает то, что *можно* сказать, но и то, что сказать *нельзя*; искусственный интеллект может генерировать все, если это *статистически вероятно*, не обращаясь к ограничениям формальной грамматики.

Примечательно, что в своих недавних исследованиях ученый, обращаясь к базовому осмыслению структур языка, на которых основаны большие языковые модели, отмечает, что сложность системы правил и структуры языка предполагает сложность задачи их воспроизведения искусственным интеллектом [3]. Научное знание переходит к синтезу, где статистическая мощь искусственного интеллекта направляется в русло структурных ограничений, описанных великим лингвистом несколько десятилетий назад.

Исследователи пытаются объединить нейросети со строгостью формальных грамматик в нейросимвольном искусственном интеллекте. Таким образом, идеи Н. Хомского определили развитие языков программирования и первых языковых моделей, теоретически они являются стандартом для измерения сложности задач (от регулярных выражений до контекстной зависимости), на практике же они реализуются созданием гибридных нейросимвольных систем.

Нейросети генерируют вероятностные распределения, предсказывая следующий пиксель (в картинках) или следующий токен (в тексте). Они создают информацию на основе усвоенных закономерностей обучающих данных. В философском смысле, по Бодрийяру [4], они генерируют гиперреальность — контент, который выглядит более упорядоченным, красивым и «правильным», чем реальный мир, но не имеет под собой физического оригинала. Большие языковые модели — и инструмент анализа, и объект изучения — обучены на больших объемах текстовых данных, они способны генерировать речь и текст, что позволяет им выполнять различные задачи, связанные с обработкой естественного языка. «В основе языковых моделей лежат нейронные сети. Они обучаются путем анализа огромных массивов текстов, выявляя закономерности, грамматические правила и смысловые связи» [5. С. 109].

Современный искусственный интеллект как «совокупность технологий компьютерных систем, имитирующих когнитивные способности человека (обучение, анализ данных, адаптация к изменяющимся условиям, принятие решений в режиме реального времени и обработка естественного языка)» [6. С. 791] не обращается к правилам грамматики: он учится на терабайтах текстов, предсказывая, вычисляя возможность следующего слова, **он** не *выводит* правила, а выражает их статистически.

В эпоху небывалого ранее прогресса нейронных сетей применение больших языковых моделей «инициирует радикальный сдвиг парадигмы в лингвистике — от эпистемологии антропоцентризма к онтологии новых „постантропных“ коммуникаций» [7. С. 145].

Классическая лингвистика ставила целью описать структуру языка (Ф. де Соссюр) или порождающие правила (Н. Хомский), современная компьютерная лингвистика в союзе с искусственным интеллектом переживает фундаментальные перемены. Если раньше (в эпоху классических алгоритмов) главной задачей было научить машину построить дерево зависимостей, то сейчас, в эпоху больших языковых моделей, задача изменилась: надо научить машину понимать смысл. Необходимо обучить языковую модель рассуждению (chain-of-thought). Это важно, чтобы перед тем, как ответить на вопрос, модель «проговаривала» про себя логические шаги, проверяла их и только после этого генерировала текст.

Лингвистические знания и методы исследования искусственного интеллекта, анализа баз данных применяются в лингвистике широко, открывая новые грани полипарадигмальных исследований: «оценка успешности моделирования отдельных элементов языковой системы и механизмов мышления, выражаемых с помощью языка, исключительно важна для систем искусственного интеллекта. Это, в свою очередь, предполагает и новые эвристические возможности для междисциплинарных исследований языка на пересечении философии, социологии, когнитивных и информационных дисциплин» [8. С. 119]. Интерес исследователей направлен на изучение лингвистического потенциала нейросетей, на анализ с позиций фреймовой семантики нарратива

[9], на возможности нейронных моделей в моделировании грамматических структур [10. С. 179]; изучаются употребление слов различных тематических групп в сгенерированной и естественной речи [11. С. 48], композиционно-структурные, семантические и пресуппозиционально-прагматические параметры сгенерированных текстов, генерация текстов в аспекте авангардной поэтики [12], исследуются способности больших языковых моделей интерпретировать референцию на материале фразеологических выражений [13], анализируется интеграция практик машинного обучения и науки о языке, оценка систем искусственного интеллекта и его эффективного обучения, где в качестве основного инструмента оценки уровня интеллектуальности системы используются языковые тесты [8. С. 119].

Междисциплинарные исследования от вопроса «Как научить машину „говорить“» перешли к изучению проблемы; «Что машина поняла о структуре языка?», «Как знание законов языка может исправить ошибки нейросетей (галлюцинации, смещения и др.)?». Лингвистика как наука, изучающая естественный человеческий язык, его структуру, функционирование, влияние на мышление, использование в обществе, нацелена на объяснение, почему язык устроен именно так, в то время как искусственный интеллект — «черный ящик» с миллиардами параметров — стремится к *предсказанию и результату*. Он выдает правильный результат, но внутренняя логика его работы абсолютно чужда человеческой логике и лингвистической теории.

«Как моделирующая наука лингвистика обладает широким спектром методов, которые наследуются другими дисциплинами, связанными с обработкой языковых данных» [7. С. 119]. Лингвисты должны стать «биологами», изучающими новый вид «организмов» — нейросети [14]. Взаимодействие лингвистики и искусственного интеллекта выражается не только в использовании текстов для обучения нейросетей, оно включает тесные неразрывные отношения связи и взаимозависимости, при которых языковые данные являются материалом, а лингвистические теории — базой для структуры алгоритмов, модифицированных языковыми практиками моделей, способных аккумулировать то, что «ускользает как от чисто технического описания (векторы, матрицы, алгоритмы), так и от традиционного философского анализа (сознание, субъект, интенциональность). В этом гибридном пространстве технические термины обретают экзистенциальную глубину, а философские концепты получают операциональную конкретность» [1. С. 319].

Развитие нейросетей как катализатор трансформации глобальной текстовой картины мира

Нейросети и алгоритмы, генерирующие текст, существовали и ранее. Тем не менее, они были очень сильно ограничены по своим функциям и могли работать только в закрытой семиотической системе с четким разделением на небольшое количество семантических категорий — например, они могли описывать данные о фондовом рынке или прогнозы погоды. Основными

чертами и даже условиями существования подобных систем были их закрытость, однотипность генерируемых текстов и узконаправленность. Искусственный интеллект способен создавать новые, оригинальные тексты [15]. Именно модели, работающие в режиме диалога, стали тем форматом, который является наиболее удобным и привычным человеку, особенно тогда, когда он не имеет опыта программирования и написания инструкций для машин. Выход ChatGPT в конце 2022 г. стал важнейшей точкой отсчета, популяризовавшей технологии машинного обучения среди пользователей, не являющихся профессионалами в области информационных технологий, и сделавшей другие модели, в том числе русскоязычные Алиса AI и GigaChat, доступными для широкого круга пользователей.

Стремительное развитие нейросетей в целом и больших языковых моделей в частности, наблюдаемое в течение последних трех лет [6] с выходом ChatGPT, а также массовое использование данных технологий полностью изменили не только возможности современных пользователей Интернета, но и послужили катализатором полной трансформации глобальной текстовой картины мира. Это связано с тем, что мы переходим от эры, когда текст был исключительно продуктом человеческого разума, к эпохе *гибридного* и *синтетического* контента. Глобальная текстовая картина мира — это все записанные знания, мнения, факты и др., доступные человечеству. Раньше этот массив данных рос органично и поэтапно, ограниченный когнитивными способностями и возможностями людей. Сейчас Интернет наполняется «шумом» — сгенерированными статьями, комментариями и отзывами, голоса людей теряются в океане синтетического текста. Нейросети обучаются на данных из интернета, поэтому, если он наводняется текстами, которые создали другие нейросети, то модели начинают обучаться на синтетических данных, что ведет к ухудшению качества, это явление носит название «самозаражения» моделей. Языковая модель усредняет реальность: исчезают диалекты, стилевое разнообразие, изобразительно-выразительные средства, культурные особенности, которые не вписываются в статистическую норму. Трансформация глобальной текстовой картины мира неизбежна. Главная проблема в том, чтобы сохранить *субъектность*. Если мы перестанем писать и думать самостоятельно, полностью делегировав это нейросетям, наша картина мира станет плоской, усредненной галлюцинацией машины.

Достижения языковых моделей в генерации текстов приводят к ложным выводам, что проблема понимания естественного языка решена. Несомненно, с позиций теоретической лингвистики современные системы демонстрируют компетентность в грамматической системе языка (*синтаксис*, *морфология*), но не могут преодолеть серьезные проблемы в *содержании* (семантика) и *намерении* (прагматика). Языковые модели владеют «формальной лингвистической компетентностью», поскольку знают правила языка, но не обладают «функциональной компетентностью», обусловленной мышлением и логикой [16].

Семиотические системы в искусственном интеллекте. Семантическая интерпретация, прагматическая контекстуальность. От обработки сигналов к интерпретации смыслов

Проблемы семиотики, семантики, прагматики и верификации истины требуют применения фундаментальной лингвистической теории в структуре машинного обучения. Интеграция семиотики в разработки искусственного интеллекта продиктована и теоретической традицией, и технологической необходимостью решения многих задач, в том числе фундаментального вопроса о том, может ли машина на самом деле понимать, или она лишь имитирует понимание через сложную статистику. Это переход от уровня «как передать данные» (сигнал) к уровню «зачем это сказано и что это значит в текущей ситуации» (смысл и цель). Семиотика, язык, коммуникация и прагматика исследуются как единое концептуальное пространство. «Семиотика по отношению к сознанию выступает как его ознакомление, означивание. Язык, в союзе с семиотикой, по отношению к сознанию выступает в качестве его репрезентанта, как средство материализации мыслительных структур [17. С. 268]. Искусственный интеллект оперирует формой, но не смыслом; статистикой, но не логикой реального мира; основные проблемы — в сфере семантики и прагматики. Кроме того, очень важны и значимы проблемы интерпретируемости («черный ящик») и этические вопросы (смещения, галлюцинации), требующие исследования. Современные нейросети отлично пользуются синтактикой, статистическим комбинированием знаков, но демонстрируют трудности с семантикой, связью знака с реальным миром и прагматикой, пониманием целей коммуникации.

Для перехода от генерации правдоподобного текста к реальному решению задач и пониманию контекста языковые модели должны научиться работать со структурой знака, различая его форму (сигнал) и содержание (смысл). Для этого нужно исследовать *рождение смысла на пересечении человеческого и искусственного интеллекта*. Следует особо подчеркнуть, что для искусственного интеллекта смысл слова — это то, как оно используется в тысячах различных контекстов, в тысячах практик [1. С. 323]. В связи с этим необходимо обратиться к вопросу о *семиотическом пределе* искусственного интеллекта, переходе от синтаксической имитации к генеративному семиозису. Активный исследовательский интерес к этим проблемам — свидетельство их возрастающей значимости. Стремительное развитие генеративного искусственного интеллекта актуализировало фундаментальные вопросы философии языка и семиотики. Если ранее машинное обучение рассматривалось преимущественно в категориях кибернетики и информатики, то появление таких мощных моделей, как GPT-4, диктует необходимость решения важного вопроса: является ли деятельность нейросети семиозисом (процессом порождения и интерпретации знаков) или же мы наблюдаем сложную форму стохастической манипуляции означающими без доступа к означаемому?

Семиотический подход к искусственному интеллекту («когнитивная семиотика» или «семиотическое моделирование») дает основание рассматривать искусственный интеллект как интерпретирующую систему: знак — объект — интерпретанта. В семиотике знак рассматривается *в единстве означающего* — форма (звук, написанное) и *означаемого* (смысл, образ предмета в реальности). Для человека «Апельсин» (означаемое) тесно связан с фруктом оранжевого цвета, сочным, цитрусовым, который можно съесть (означаемое). Для языковой модели токен [Апельсин] — *это вектор, набор цифр* в многомерном пространстве, имеющий математическую близость к векторам [Фрукт], [Оранжевый], [Вкусный], но он не связан с реальным апельсином, что свидетельствует о разрыве означающего и означаемого. Языковая модель руководствуется означающими, она оперирует «оболочками» слов, создавая своего рода орнаменты и узоры, *предсказывая возможность и вероятность следующего токена на основе предыдущих контекстов*.

Закономерно возникает вопрос: есть ли в сгенерированном тексте отношение к реальности или мы наблюдаем только взаимодействие знаков? Жан Бодрийяр называл это «симулякром третьего порядка» — знаком, который маскирует отсутствие глубинной реальности, перенеся исследования симулякра из сферы онтологии в описание современной социальной реальности, описывая процесс постепенного исчезновения реальности и замены ее знаками, утверждая, что «симулякр — это вовсе не то, что скрывает собой истину, — это истина, скрывающая, что ее нет» [4. С. 170]. Когда нейросеть пишет эссе о чувствах, которых у нее нет, или описывает историческое событие, которого не было, она делает вид, что за этими словами стоит опыт или факт. Но за ними — только математическая вероятность. «С точки зрения семиотического направления, «симулякр» — самореференциальный знак (образ несуществующей действительности), подменяющий символ, который олицетворяет слово, и его конвенциональные значения» [18. С. 146].

У модели нет доступа к физическому миру, но она *создает такую детальную систему языковых вероятностей*, что демонстрирует свойства, похожие на логическое рассуждение. Современные модели, которые обучаются на триллионах параметров, можно рассматривать как реализацию такого процесса, ведь они не познают мир (физическую реальность), они познают *тексты о мире* (знаковую реальность). «ИИ-симулякры — это не просто обманчивые факсимиле; они становятся действующими реальностями внутри цифровой культуры, все больше формируя дискурс, эстетику и даже эмоциональные реакции в отсутствие человеческой субъектности» [19]. Текст искусственного интеллекта — это копия человеческого мышления, но за ней не стоит думающий субъект, вероятность представлена вместо истины.

Проблема генеративного искусственного интеллекта — отсутствие референта, объекта в реальном мире, на который указывает слово. Для человека язык *интенционален* (нацелен на предмет, мы говорим о чем-то), а для модели — нет, она предсказывает следующий токен на основе предыдущего.

В языковых моделях это проявляется следующим образом: обучающая выборка (Dataset) — это отражение цифрового пространства, которое само по себе является набором репрезентаций, мнений и копий, нейросеть учится не на реальности, а на ее отражении, как результат искусственный интеллект генерирует *копию копии без оригинала*. Это объясняет феномен «галлюцинаций» с технической точки зрения: галлюцинация — это ошибка, но по Ж. Бодрийяру, — это торжество симулякра. Галлюцинация не является ложью, потому что у искусственного интеллекта нет концепции «истины», это чистый знак, освобожденный от обязанности что-либо означать. «Инструменты генеративного ИИ создают тексты, изображения и аудиовизуальный контент, которые не отсылают к какой-либо единственной реальности или авторскому замыслу, а скорее возникают из вероятностных моделей, обученных на огромных корпусах ранее существовавших данных. Этот процесс воплощает чистый симулякр Бодрийяра, в котором результат не является ни копией, ни подделкой, а совершенно новой конструкцией, оторванной от первоисточника» [20].

Исследования показывают: если обучать новые модели на данных, сгенерированных предыдущими моделями, качество деградирует. Симулякры начинают копировать симулякры. Реальность (человеческий опыт) вымывается из обучающей выборки. Коммуникация, по Бодрийяру, требует обмена, но ведь диалог с искусственным интеллектом — не обмен, а *бесконечное отражение запросов* пользователей. Эффективное статистическое моделирование языка выявляет и идентифицирует модели и образцы в огромных массивах данных, но языковая модель не понимает физической реальности и причинно-следственных связей, у нее нет тела и опыта взаимодействия с окружающей действительностью [21].

Одна из фундаментальных проблем в области искусственного интеллекта — проблема *семантики*, или способности машин извлекать, интерпретировать и генерировать смысл: можно идеально манипулировать символами по правилам, не понимая, что они значат. Несмотря на достижения в создании больших языковых моделей, увеличение их количества, объема и мощности, разработку более совершенных методов глубокого восприятия и интерпретации контекста [5. С. 110], выделяются существенные проблемы, связанные с их пониманием и обработкой языка. Искусственный интеллект сейчас находится на этапе «*семантической имитации*». Он уверенно предсказывает, какое слово *должно* идти следующим, но не *понимает*, почему оно там стоит.

Вопрос о том, может ли машина «понимать» *смысл* обрабатываемой информации, является одним из самых важных. В эпоху доминирования больших языковых моделей граница между *синтаксической обработкой данных* и *семантическим осмыслением* стала еще более размытой. Переход от статистической обработки текста к реальному смысловому пониманию — главная задача на пути к совершенствованию искусственного интеллекта. Главной

проблемой компьютерной лингвистики остается разрыв между статистической вероятностью появления слова и его смысловым наполнением. Языковые модели обучаются предсказывать следующий токен на основе массивов данных. Они «знают», какие слова часто стоят рядом, но не имеют когнитивной модели мира, стоящей за этими словами, не понимают, «как функционирует реальный мир. Они могут распознавать закономерности, но не понимают причинно-следственных связей, <...> ориентированы на выявление статистических закономерностей в данных, чем на понимание их смысла» [5. С. 110]. В области *семантики, проблем смысла и значения* единиц языка и знаковых систем, искусственный интеллект использует векторные представления, слова и фразы преобразуются в математические векторы. *Близость векторов означает близость значений*. В семантике значение описывается через *семы (компоненты смысла)*, в языковых моделях оно описывается через *координаты вектора*. Несмотря на то, что искусственный интеллект предсказывает токен, формируя цепочки размышлений, распределяет вычисление в зависимости от сложности задачи [6. С. 792], трудные случаи полисемии ему недоступны. Например, фраза «*Предприятие стоит*» без расширенного контекста недоступна в полной мере искусственному интеллекту: означает ли это, что предприятие не работает (бездействие), или что оно находится в определенном месте (локация), или речь идет о стоимости. Слова имеют разные значения в зависимости от контекста. Для человека выбор значения происходит автоматически благодаря жизненному опыту, а для искусственного интеллекта — это математическая вероятность.

Семантическая энтропия как фундаментальное свойство языка.

Энтропия как мера непредсказуемости символа

Энтропия (*греч.* *Εντροπία* — *поворот, превращение*, мера беспорядочности или разупорядоченности системы) как универсальное естественнонаучное понятие прочно вошло в гуманитарные и компьютерные науки. Специалист по языковым моделям видит в ней меру неопределенности некоторой ситуации [22], «связанной с данными: чем более неопределенны данные, тем выше энтропия. Точнее, энтропия связана с количеством информации, содержащейся в данных, то есть чем более неопределенны данные, тем больше информации требуется, чтобы описать их» [23. С. 153]. Лингвист видит в энтропии возможности для творчества, определяющие потенциал языка [24]. Различные подходы к трактовке энтропии открывают разнообразные возможности исследования этого феномена. Уточним, что в русской лингвистической традиции энтропия — не «шум», а «глубина», не мера непредсказуемости символа, а явление, имеющее глубокие корни, уходящие в структурную лингвистику, семиотику и теорию информации.

Семантическая энтропия рассматривается как фундаментальное свойство языка, обеспечивающее его гибкость и способность передавать новые

смыслы; это мера неопределенности значения языкового знака (слова, фразы) вне контекста, важная характеристика словесного творчества.

Интересно в этой связи обратиться к *избыточности*, *предсказуемости*, *информативности* и *новизне*. Низкая энтропия (высокая упорядоченность/предсказуемость) и высокая избыточность (наличие элементов, дублирующих информацию) тесно связаны в языке, что вполне объяснимо, поскольку чем выше избыточность, тем ниже его энтропия, так как следующее слово или букву легче предсказать. Эти особенности важны в таких речевых ситуациях, где достоверность информации более приоритетна, чем ее оригинальность и новизна. К ним относятся этикетные ритуальные фразы, документы, протоколы, инструкции, в которых используются клише, сложные синтаксические конструкции и повторы, чтобы избежать разночтений, обеспечивающих однозначное понимание. Рифма, ритм и устойчивые эпитеты понижают энтропию, ограничивают выбор слова, а во фразеологизмах, воспринимающихся как единое целое, энтропия стремится к нулю. *Низкая энтропия* и *высокая избыточность* — фундамент и основа языка, обеспечивающие быстрое понимание, в том числе и при беглом чтении (мы не читаем все буквы, мы угадываем слова). Если текст с низкой энтропией можно читать поверхностно, то текст с высокой энтропией требует внимания и сосредоточенности, потому что, пропустив одно слово, мы можем потерять смысл всего сообщения.

Высокая энтропия в языке — непредсказуемость следующего слова на основе предыдущих, информационная плотность и *новизна* сообщений, в которых очень мало избыточности, где каждое слово несет информационную нагрузку. Это язык творчества, нарушение стандартных связей слов с невозможностью достоверно предсказать следующее слово, что требует усилий для расшифровки, но рождает новые смыслы. Чем больше у слова значений (полисемия), тем выше его семантическая энтропия; таким образом, процесс понимания сообщения — процесс *подавления энтропии* с помощью контекста, где из веера возможных значений выбирается правильное. «Чем выше энтропия, тем больше неопределенности или многозначности содержится в сообщении, что делает его менее предсказуемым или однозначным» [22. С. 80]. Например, низкая энтропия характеризует канцелярский язык или разнообразные команды и указания.

Если бы слова имели четко фиксированные значения, как в языке программирования, эволюция мысли остановилась бы, а человек не смог бы выразить ничего нового, только комбинировать старое. Смыслы не жестко привязаны к знакам, они размыты, что дает возможность людям понимать друг друга не как машины, обменивающиеся командами, а как личности, «соприкасающиеся мирами». Значение слова — не точка, а функция распределения вероятностей [25]. Высокая энтропия рассматривается как *ценность*. Ю.М. Лотман отмечал, что культура стремится повышать свою энтропию, чтобы не застыть. В семиотике «ошибка» (нарушение нормы) — это часто рождение нового значения.

В классической теории информации (К. Шеннон) энтропия — *мера неопределенности или «свободы выбора»*. Для понимания энтропии как возможности нужно опираться не на смысл сообщения в его лингвистическом осмыслении, а на *вероятность* его появления. «Энтропия — это мера неопределенности содержания текста, она связана с объемом информации, заключенной в данных: чем более они неопределенны, тем больше информации требуется для их описания» [22. С. 78]. Необходимо подчеркнуть, что понятие энтропии трансформируется при переходе от теории информации к системам искусственного интеллекта, и поэтому для нейросети и поэта «хаос» имеет разное значение; энтропия = неопределенность = объем информации является не метафорой, а строгим математическим законом.

Неопределенность — это количество вариантов, из которых можно выбрать, информация — результат выбора. Чем он сложнее (чем выше была неопределенность), тем важнее результат. Объем описания — цена, которую мы платим за сохранение уникальности этого выбора. Хаос требует детального описания каждого элемента, в то время как порядок позволяет использовать краткие формулы. Текст с высокой энтропией информационно насыщен и требует большого объема памяти для хранения, так как он непредсказуем. Для современных больших языковых моделей энтропия — это рабочий инструмент: искусственный интеллект стремится снизить неожиданность следующего слова. Идеальная модель — та, для которой текст из обучающей выборки имеет минимальную энтропию (она точно знает, что будет дальше). Для искусственного интеллекта энтропия — это ландшафт, структурирование и символизирование окружающей среды, это математическая задача. Для семиотики энтропия — это пространство, где рождаются искусство, юмор и новые культурные коды. Понимание этих двух граней энтропии необходимо для создания искусственного интеллекта нового поколения, который сможет не просто генерировать статистически верный текст, но и понимать природу творчества. Переход от энтропии токенов к семантической энтропии знаменует собой сдвиг от анализа *синтаксиса* к анализу *семантики* в искусственном интеллекте. Семантическая энтропия переносит измерение неопределенности из пространства лексики (слова/символы) в пространство смыслов (семантика). Переход от лексического пространства к семантическому — фундаментальный сдвиг в том, как мы оцениваем «уверенность» искусственного интеллекта. Семантическая энтропия — мост между математической теорией информации и гуманитарной наукой о смыслах. Этот инструмент критически важен не только для теоретической лингвистики, но и для создания надежного искусственного интеллекта, способного понимать, когда он говорит правду, а когда — генерирует шум.

Прагматика и контекстуальная «слепота»

Серьезная проблема искусственного интеллекта — неспособность понимать контекст, намерения и скрытые смыслы, очевидно, что между

«правильным» текстом модели и «осмысленным» общением лежит огромная пропасть. Отсутствие прагматической составляющей — главная причина, по которой языковые модели совершают ошибки в простых ситуациях. Прагматика изучает отношения субъекта коммуникации к языку и экстралингвистический контекст, *значение предложения* (буквальный смысл) и *значение говорящего* (что человек хотел сказать). Искусственный интеллект не всегда хорошо справляется с распознаванием иронии, сарказма и культурных аллюзий, так как они не содержатся в самом сигнале (тексте), а находятся в надсистеме (в культуре, ситуации и др.).

Сарказм часто строится на противоречии между позитивной лексикой и негативным контекстом, или тональностью. В моделях нет просодической информации (интонация, ударение), визуальных указаний и помощи (мимика, жесты), поэтому они воспринимают текст буквально. Научить машину грамматике сложно, но возможно, так как грамматика логична. Научить машину прагматике — значит научить ее понимать саму суть человеческого существования: нашу нелогичность, наши недомолвки, нашу культуру и наш юмор.

Язык пронизан аллюзиями и культурными отсылками. Утверждение «Он настоящий Плюшкин» для носителя русского языка означает патологическую скупость человека, страсть к накопительству и деградацию личности. Для искусственного интеллекта, обученного преимущественно на англоязычных текстах или технической документации, эта метафора может быть потеряна или переведена буквально, если в обучающей выборке не было книг Н.В. Гоголя с комментариями.

В силу того, каким образом обучаются нейросети, следует отметить англоцентризм многих больших языковых моделей, которые изначально обучались на английском языке, на котором представлено наибольшее количество индексируемой информации в Интернете. Это приводит к наличию большого количества англицизмов на всех уровнях языка, вплоть до таких очевидных аспектов, как отсутствие согласования между членами предложения. Эта проблема постепенно решается как создателями глобальных моделей искусственного интеллекта, совершенствующих разметку лексических массивов данных, так и русскоязычными компаниями. Помимо этого повсеместное использование больших языковых моделей ведет к унификации и стандартизации речи, а также сопровождается искоренением региональной вариативности и лексико-грамматических особенностей.

Ситуации с растущим недоверием к текстам и информации, сгенерированной нейросетями, может быть противопоставлено чрезмерное доверие к подобным технологиям без дальнейшей верификации, что может привести к серьезным проблемам в различных профессиональных сферах. Одним из первых известных примеров подобного использования может служить ситуация, возникшая в июне 2023 г. — спустя полгода после выхода ChatGPT [26]. Адвокат, использовавший раннюю модель ChatGPT, не имеющую выхода в

Интернет, сгенерировал юридическое заключение, которое затем было представлено в суд Нью-Йорка. Документ содержал шесть ссылок на судебные прецеденты, которые выглядели правдоподобно, имели непосредственное отношение к рассматриваемому делу и были грамотно составлены, но при этом не существовали вообще. Когда непреднамеренная фальсификация была раскрыта, автор текста заявил, что не осознавал, что искусственный интеллект не ищет прецеденты в базе данных, а лишь создает текстовые объекты, похожие по структуре на то, что заложено в его базу знаний.

Ситуация улучшается, так как подключаются модули поиска по Интернету, а в новых моделях появился режим размышления, который позволяет искусственному интеллекту многократно проанализировать запрос пользователя и написать подробные инструкции для себя, «переведя» его с естественного языка, производимого человеком, в упрощенную форму данного естественного языка, а затем — в машинный код. Этот алгоритм открывает уникальную с точки зрения коммуникации ситуацию саморефлексии машины.

Алгоритмы обнаружения текстов, написанных большими языковыми моделями. Перплексия как оценка качества языковых моделей

Несмотря на новейшие разработки и большой прогресс, проблема создания текстов, основанных на несуществующих фактах, ни в коем случае не является решенной. Такой феномен носит название «галлюцинации», то есть уверенной реакции системы искусственного интеллекта, не соответствующей исходным данным. Несмотря на высокий уровень развития систем и даже на новейшие способности «размышления», с эпистемологической точки зрения система не в состоянии провести подробный самоанализ и осознать отношение своих высказываний к действительности.

Упоминая эпистемологические особенности, следует также обсудить так называемый эффект когнитивной амнезии Гелл-Манна — феномен, описывающий парадоксальное доверие людей к средствам массовой информации. Читатели положительно или нейтрально воспринимают информацию, которая находится за пределами их области экспертизы, но после того, как сталкиваются с хорошо знакомой темой, видят поверхностность предоставляемой информации и могут начать скептически относиться не только к рассматриваемому материалу, но и его источнику в целом.

Распространение генеративного искусственного интеллекта повышает важность человеческого авторитета и доверия к текстам, написанным человеком при условии последующей верификации их аутентичности.

Чтобы понять, каким образом работают алгоритмы обнаружения текстов, написанных большими языковыми моделями, следует понять принцип работы GPT, то есть генеративных предобученных трансформеров. Принцип их работы схож с механизмами предиктивного ввода, который существовал еще

несколько десятилетий назад на различных устройствах. Он может быть основан на завершении отдельной лексемы или словосочетания. Большие языковые модели, в свою очередь, работают с более крупными единицами языка и имеют более объемный контекст — у современных моделей он может превышать миллионы знаков.

В таких случаях вывод слов, словосочетаний, предложений и целых текстов строится в первую очередь на статистической вероятности, которая построена на основе результата обучения языковой модели. Так как современные неспециализированные модели, используемые пользователями интернета, содержат крайне объемный массив данных, даже сами разработчики программного обеспечения, создающие их, часто не могут предсказать их возможности. Тем не менее следующая единица, выдаваемая большой языковой моделью, предсказывается статистически, как и в языковых корпусах, основываясь на частотности сочетаний.

Для демонстрации и введения понятия перплексии рассмотрим пример сочетаемости лексемы «добрый» взятый из Национального корпуса русского языка воспользовавшись инструментом n-gram, где n — число лексических единиц. На примере ниже можно отметить, что сочетание «доброе утро» является гораздо более частотным, чем «доброе отношение». Это значит, что вероятность получить данное слово при генерации текста при равнозначных остальных параметрах гораздо выше, и большая языковая модель имеет тенденцию придерживаться определенных слов и выражений (табл. 1).

Сочетаемость лексемы «добрый» / Collocation of the lexeme “good”

№	Контекстов / Contexts	IPM	Отрывок Examples
1	1,112	2.86	доброе дело
2	832	2.14	доброе утро
3	558	1.43	доброе сердце
4	415	1.07	доброе слово
5	409	1.05	доброе имя
6	231	0.59	доброе лицо
7	188	0.48	доброе старое
8	149	0.38	доброе отношение

Источник: исследование М.Л. Новиковой и Ф.Н. Новикова.

Source: compiled by Narina L. Novikova & Philipp N. Novikov.

Основной параметр, позволяющий определить предсказуемость текста с помощью машинных средств, носит название перплексии. Перплексия является статистическим критерием и рассчитывается по следующей формуле:

$$PP(h) := 2^{H(p)} = 2^{-\sum_x p(x) \log_2 p(x)},$$

где $PP(p)$ — перплексия, $H(p)$ — энтропия, которая в данном случае обозначает хаотичность, неопределенность. Высокая степень перплексии

соответствует неожиданности и отклонению от статистической нормы, а низкая показывает высокую предсказуемость. Таким образом, в тексте с высокой перплексией языковая модель может с более высокой вероятностью угадать следующий токен (языковой знак, длина которого в зависимости от языка составляет определенное количество символов или слов), основываясь на данных о частотности, подобных тем, которые были представлены выше в примере из корпуса русского языка.

Важно также отметить, что в больших языковых моделях случайность регулируется таким показателем, как «температура». Этот параметр обычно варьируется от 0 до 1, где 0 — максимально детерминистическая ситуация, а 1 — наиболее произвольная, но при этом возможно, более релевантная определенному запросу ситуация. При параметре «температура», равном нулю, моделью должен выдаваться один и тот же ответ при одинаковом вводе данных.

Механизмы распознавания использования генеративного искусственного интеллекта полагаются в первую очередь именно на параметр перплексии. Это означает, что алгоритму нужно сначала предсказать наиболее вероятные сценарии генерации текста, а затем сопоставить их с тем текстом, который представлен в анализируемом документе. Таким образом, механизм детекции использования большой языковой модели сам по себе является большой языковой моделью. С другой стороны, при высоком параметре температуры повышается вероятность получения галлюцинаций, что с точки зрения семантики облегчает задачу по обнаружению сгенерированных текстов людьми, но не алгоритмами, которым становится сложнее найти лексико-семантические закономерности. Исследование, проведенное на основе массива данных, в который были включены аннотации к научным статьям, показало, что перплексия написанных человеком аннотаций к научным статьям с медианой 35,9 значительно выше, чем у тех, которые были сгенерированы искусственным интеллектом [27].

Помимо понятия перплексии и общего уровня предсказуемости текстов одним из надежных показателей использования генеративного искусственного интеллекта является всплеск популярности следующих лексических единиц, многие из которых, как считается, появились под влиянием большого коллектива людей, владеющих определенным региональным вариантом английского языка и занимавшихся первоначальной разметкой данных для нейросети ChatGPT: «delve into», «in summary», «take a dive into», «navigating the landscape» и т.д. Таким образом, определенные клише и устойчивые выражения, которые хотя и являются естественными для человеческой речи, выдаются искусственным интеллектом с гораздо более высокой стабильностью.

Среди синтаксических особенностей, которые проявляются в текстах, сгенерированных нейросетями, можно отметить наличие «Em dash» (длинного тире) — символа, являющегося одним из самых известных примеров автоматически сгенерированных текстов. Этот символ является нехарактерным

для большого количества письменных стилей и расположен во многих раскладках на клавиатурах в неудобных местах и является одним из самых явных признаков использования генеративного искусственного интеллекта.

Использование различных эмодзи, то есть графических символов, встроенных в современные операционные системы, а также списков с буллитными знаками также можно отнести к явным признакам использования искусственного интеллекта. Если каждый из них по отдельности не всегда является явным признаком того, что автор прибегал к генерации текста, то совокупность данных признаков приводит к ощутимому кумулятивному эффекту со стороны читателей, позволяющего распознавать произведения больших языковых моделей.

Повышенное внимание к определенным структурным признакам текста со стороны читателей может привести к обратному эффекту, при котором они могут подозревать авторов в использовании нейросетей для генерации текста, если обнаруживают хотя бы один из вышеупомянутых признаков, например — длинное тире. Ложное срабатывание механизмов присуще также и автоматическим системам опознавания нейросетевых текстов. Подобная ситуация может сложиться тогда, когда в систему загружается текст, данные о структуре которого заранее заложены в нее, или тогда, когда текст, написанный человеком, имеет определенные структурные признаки машинного текста, но при этом написан самостоятельно.

Заключение

Научные изыскания находятся в точке фазового перехода искусственный интеллект перерос стадию «стохастических попугаев» (чистая обработка сигналов), но еще не достиг стадии «субъекта коммуникации» (полноценная прагматика). Проблема заключается в том, что мы пытаемся извлечь человеческий смысл из математической функции. И пока неясно, может ли количество (вычислительная мощность и данные) перейти в качество (понимание интенций), или для этого нужно принципиально иное, не вероятностное построение, а когнитивное.

Сопряжение лингвистики и искусственного интеллекта переходит на новый этап. Период «чистой инженерии», когда казалось, что достаточно лишь увеличить объем данных и мощность серверов, заканчивается. Будущее — за гибридными, полипарадигмальными системами, где статистическая мощь нейросетей направляется и контролируется структурированными знаниями о языке и мире. Это позволит создать искусственный интеллект, который не просто имитирует владение языком, но и понимает его коммуникативную природу. Идеального алгоритма обнаружения искусственного интеллекта на данный момент не существует. Статистические методы работают до первой серьезной правки текста человеком, поэтому человеческий фактор до сих пор является крайне важным.

Список литературы

1. *Аришинов В.И., Янукович М.Ф., Гарашко Ю.Н.* Онтология в диалоге: рождение языка и смысла на пересечении человеческого и искусственного интеллекта // Вестник Пермского университета. Философия. Психология. Социология. 2025. № 3. С. 317–328. <https://doi.org/10.17072/2078-7898/2025-3-317-328> EDN: AYNKYE
2. *Бедненко Ю.И.* Искусственный интеллект и проблема сознания: философские и лингвистические аспекты // Гуманитарные исследования. История и филология. 2025. № 17. С. 103–113. <https://doi.org/10.24412/2713-0231-2025-17-103-113> EDN: LANAJG
3. *Бодрийяр Ж.* Симулякры и симуляция. Тула, 2013.
4. *Гусаренко С.В., Гусаренко М.К.* Композиционно-структурные, семантические и пресуппозиционально-прагматические параметры и дефекты сгенерированных коротких текстов в языковой нейросети GigaChat // Гуманитарные и юридические исследования. 2024. Т. 11. № 2. С. 368–379. <https://doi.org/10.37493/2409-1030.2024.2.21> EDN: JMESZZ
5. *Гусаренко С.В., Гусаренко М.К.* Предикатно-фреймовое описание линейной семантики художественного нарратива // Филологические науки. 2020. № 3. С. 18–30. <https://doi.org/10.20339/PhS.3-20.018> EDN: XZUUXH
6. *Демченков С.А., Федяева Н.Д.* Вербальный хеппенинг: о некоторых моделях компьютерной генерации текстов в аспекте авангардной поэтики // Humanitarian Vector. 2020. Т. 15. № 1. С. 27–37. <https://doi.org/10.21209/1996-7853-2020-15-1-27-37> EDN: ZJTBAV
7. *Зырянова И.Н.* Генеративные языковые модели и феномен антиантропоцентризма — новые перспективы лингвистической парадигмы «posthumano» и «общего/сильного» // Известия Байкальского государственного университета. 2024. Т. 34. № 1. С. 144–152. [https://doi.org/10.17150/2500-2759.2024.34\(1\).144-152](https://doi.org/10.17150/2500-2759.2024.34(1).144-152) EDN: NGYEЕB
8. *Матурана У., Варела Ф.* Дерево познания: биологические корни человеческого понимания. М. : Прогресс-Традиция, 2001.
9. *Назаров Н.В., Шульгинов В.А.* «Первый блин комом»: ограничения и потенциал LLM в распознавании референции устойчивых фраз // Коммуникативные исследования. 2024. Т. 12. № 4. С. 881–896. [https://doi.org/10.24147/2413-6182.2025.12\(4\).881-896](https://doi.org/10.24147/2413-6182.2025.12(4).881-896) EDN: ZBGBZR
10. *Налимов В.В.* Вероятностная модель языка. О соотношении естественных и искусственных языков. М. : Наука, 1974. EDN: QXGCTG
11. *Саакян Л.Н.* ИИ как гуманизированный субъект: митигативные стратегии искусственного интеллекта // Коммуникативные исследования. 2025. Т. 12. № 4. С. 789–804. [https://doi.org/10.24147/2413-6182.2025.12\(4\).789-804](https://doi.org/10.24147/2413-6182.2025.12(4).789-804) EDN: CJFFSS
12. *Студеникина К.А.* Оценка языковой способности нейронных моделей на материале предикативного согласования в русском языке // Труды Института системного программирования РАН. 2022. № 34 (6). С. 179–184. [https://doi.org/10.15514/ispras-2022-34\(6\)-14](https://doi.org/10.15514/ispras-2022-34(6)-14) EDN: UCJBEV
13. *Тельнов Р.Е., Ларцина С.В.* Типовые различия естественных и сгенерированных нейронной сетью текстов в количественном аспекте // Научный диалог. 2023. Т. 12. № 7. С. 47–65. <https://doi.org/10.24224/2227-1295-2023-12-7-47-65> EDN: ZZRRKN
14. *Хомский Н.* Синтаксические структуры // Новое в лингвистике. М., 1962. Вып. II. С. 412–527.
15. *Шаврина Т.О.* О методах компьютерной лингвистики в оценке систем искусственного интеллекта // Вопросы языкознания. 2021. № 6. С. 117–138. <https://doi.org/10.31857/0373-658X.2021.6.117-138> EDN: PFGVQJ
16. *Шарандин А.Л.* Сознание, семиотика, язык, коммуникация и прагматика как единое концептуальное пространство // Вопросы психолингвистики. 2015. № 3 (25). С. 240–250. EDN: UDLHJT
17. *Шевченко Е.В.* Симулякр как тактический прием оказания ориентирующего воздействия // Вопросы когнитивной лингвистики. 2015. № 3 (44). С. 144–150. EDN: TTYPMF

18. Павловская О.Е., Ильина Л.Е. Энтропия текста как рассеивание смысла // Вестник Северного (Арктического) федерального университета. Серия: Гуманитарные и социальные науки. 2023. Т. 23. № 6. С. 77–86. <https://doi.org/10.37482/2687-1505-V308> EDN: ZPZZPW
19. Шелестюк Е.В., Щетинкина Е.А. Стохастичность и энтропия в лингвистике // Вестник Челябинского государственного университета. 2023. № 2 (472). С. 150–165. EDN: WIVZGY
20. Андриенко Т.П. Энтропия текста как предпосылка реализации стратегии перевода // Известия Гомельского государственного университета им. Ф. Скорины. 2017. № 1 (100). С. 76–80. EDN: YFYPLF
21. Baroni M. On the proper role of linguistically-oriented deep net analysis in linguistic theorizing // Algebraic Approaches to Sequence Data. 2022. <https://doi.org/10.1201/9781003205388-1>
22. Chomsky N., Roberts I., Watumull J. Noam Chomsky: The False Promise of ChatGPT // The New York Times. 2023. 8 March.
23. Elek A., Yildiz H.S., Akca B., Oren N.C., Gundogdu B. Evaluating the efficacy of perplexity scores in distinguishing AI-generated and human-written abstracts // Academic Radiology. 2025. Vol. 32. Iss. 4. P. 1785–1790. <https://doi.org/10.1016/j.acra.2025.01.017> EDN: HMLYED
24. Jiang S., Li K. Another kind of authenticity: the visual simulacra of artificial intelligence // Digital Creativity. 2025. Vol. 36. Iss. 1. P. 75–91. <https://doi.org/10.1080/14626268.2025.2467087> EDN: ZRIUEB
25. Konzack L. Generative AI, Simulacra, and the Transformation of Media Production // Athens Journal of Mass Media & Communications. 2025. Vol. 11. Iss. 3. <https://doi.org/10.30958/ajmmc.11-3-3> EDN: BYIUCE
26. Lyon C.F. Fake cases, real consequences: misuse of ChatGPT leads to sanctions // NY Litigator. 2023. Vol. 28. Iss. 2. P. 8–12.
27. Mahowald K., Ivanova A.A. et al. Dissociating language and thought in large language models: a cognitive perspective // Trends in Cognitive Sciences. 2024. <https://doi.org/10.1016/j.tics.2024.01.011> EDN: SKTATZ
28. Miralles-González P., Huertas-Tato J., Martín A., Camacho D. Not all tokens are created equal: Perplexity Attention Weighted Networks for AI generated text detection // arXiv. 2025. <https://doi.org/arXivpreprint arXiv:2501.03940>
29. Pudasaini S., Miralles-Pechuán L., Lillis D., Llorens Salvador M. Survey on AI-generated plagiarism detection: The impact of large language models on academic integrity // Journal of Academic Ethics. 2024. <https://doi.org/10.48550/arXiv.2407.13105>
30. Ram A. Generative AI and Text Creation. Preprint. 2023. <https://doi.org/10.48550/arXiv.2407.13105>

References

1. Arshinov, V.I., Yanukovich, M.F., & Garashko, Yu.N. (2025). Ontology in Dialogue: Birth of Language and Meaning at the Intersection of Human and Artificial Intelligence. *Perm University Herald. Philosophy. Psychology. Sociology*, 3, 317–328. (In Russ.). <https://doi.org/10.17072/2078-7898/2025-3-317-328> EDN: AYNKYE
2. Bednenko, Yu.I. (2025). Artificial Intelligence and the Problem of Consciousness: Philosophical and Linguistic Aspects. *Humanitarian Studies. History and Philology*, 17, 103–113. (In Russ.). <https://doi.org/10.24412/2713-0231-2025-17-103-113> EDN: LANAJG
3. Baudrillard, J. (2013). *Simulacra and Simulation*. Tula. (In Russ.).
4. Gusarenko, S.V., & Gusarenko, M.K. (2024). Compositional-Structural, Semantic, and Presuppositional-Pragmatic Parameters and Defects of Generated Short Texts in the Gigachat Language Neural Network. *Humanities and Law Studies*, 11(2), 368–379. (In Russ.). <https://doi.org/10.37493/2409-1030.2024.2.21> EDN: JMESZZ

5. Gusarenko, S.V., & Gusarenko, M.K. (2020). Predicate-Frame Description of the Linear Semantics of a Literary Narrative. *Philological Sciences*, 3, 18–30. (In Russ.). <https://doi.org/10.20339/PhS.3-20.018> EDN: XZUHX
6. Demchenkov, S.A., & Fedyaeva, N.D. (2020). Verbal Happening: On Some Models of Computer Text Generation in the Aspect of Avant-Garde Poetics. *Humanitarian Vector*, 15(1), 27–37. (In Russ.). <https://doi.org/10.21209/1996-7853-2020-15-1-27-37> EDN: ZJTBAV
7. Zyryanova, I.N. (2024). Generative Language Models and the Phenomenon of Antianthropocentrism: New Perspectives of the Linguistic Paradigm “Posthumano” and “General/Strong” AI. *Bulletin of Baikal State University*, 34(1), 144–152. (In Russ.). [https://doi.org/10.17150/2500-2759.2024.34\(1\).144-152](https://doi.org/10.17150/2500-2759.2024.34(1).144-152) EDN: NGYEEB
8. Maturana, H., & Varela, F. (2001). *The Tree of Knowledge: The Biological Roots of Human Understanding*. Moscow: Progress-Traditsiya. (In Russ.).
9. Nazarov, N.V., & Shulginov, V.A. (2024). “The First Pancake is Always Lumpy”: Limitations and Potential of LLM in Recognizing the Reference of Set Phrases. *Communication Studies*, 12(4), 881–896. (In Russ.). [https://doi.org/10.24147/2413-6182.2025.12\(4\).881-896](https://doi.org/10.24147/2413-6182.2025.12(4).881-896) EDN: ZBGBZR
10. Nalimov, V.V. (1974). *Probabilistic Model of Language. On the Relationship between Natural and Artificial Languages*. Moscow: Nauka. (In Russ.). EDN: QXGCTG
11. Saakyan, L.N. (2025). AI as a Humanized Subject: Mitigative Strategies of Artificial Intelligence. *Communication Studies*, 12(4), 789–804. (In Russ.). [https://doi.org/10.24147/2413-6182.2025.12\(4\).789-804](https://doi.org/10.24147/2413-6182.2025.12(4).789-804) EDN: CJFFSS
12. Studenikina, K.A. (2022). Evaluation of the Language Ability of Neural Models Based on Predicative Agreement in Russian. *Proceedings of the Institute for System Programming of the RAS*, 34(6), 179–184. (In Russ.). [https://doi.org/10.15514/ispras-2022-34\(6\)-14](https://doi.org/10.15514/ispras-2022-34(6)-14) EDN: UCJBEV
13. Telpov, R.E., & Lartsina, S.V. (2023). Typical Differences between Natural and Neural Network-Generated Texts in a Quantitative Aspect. *Scientific Dialogue*, 12(7), 47–65. (In Russ.). <https://doi.org/10.24224/2227-1295-2023-12-7-47-65> EDN: ZZRRKN
14. Chomsky, N. (1962). Syntactic Structures. In: *New in Linguistics* (Vol. 2, pp. 412–527). Moscow. (In Russ.).
15. Shavrina, T.O. (2021). On the Methods of Computational Linguistics in Evaluating Artificial Intelligence Systems. *Voprosy Jazykoznanija*, 6, 117–138. (In Russ.). <https://doi.org/10.31857/0373-658X.2021.6.117-138> EDN: PFGVQJ
16. Sharandin, A.L. (2015). Consciousness, Semiotics, Language, Communication and Pragmatics as a Single Conceptual Space. *Journal of Psycholinguistics*, 3(25), 240–250. (In Russ.). EDN: UDLHJT
17. Shevchenko, E.V. (2015). Simulacrum as a Tactical Method of Rendering an Orienting Influence. *Issues of Cognitive Linguistics*, 3(44), 144–150. (In Russ.). EDN: TTYPMF
18. Pavlovskaya, O.E., & Ilyina, L.E. (2023). Text Entropy as Meaning Dispersion. *Vestnik of Northern (Arctic) Federal University. Series: Humanitarian and Social Sciences*, 23(6), 77–86. (In Russ.). <https://doi.org/10.37482/2687-1505-V308> EDN: ZPZZPW
19. Shelestyuk, E.V., & Shchetinkina, E.A. (2023). Stochasticity and Entropy in Linguistics. *Bulletin of Chelyabinsk State University*, 2, 150–165. (In Russ.). EDN: WIVZGY
20. Andrienko, T.P. (2017). Text Entropy as a Prerequisite for the Implementation of Translation Strategy. *Proceedings of Francysk Skorina Gomel State University*, 1 (100), 76–80. (In Russ.). EDN: YFYPLF
21. Baroni, M. (2022). On the Proper Role of Linguistically-Oriented Deep Net Analysis in Linguistic Theorizing. *Algebraic Approaches to Sequence Data*. <https://doi.org/10.1201/9781003205388-1>
22. Chomsky, N., Roberts, I., & Watumull, J. (2023). Noam Chomsky: The False Promise of ChatGPT. *The New York Times*, 8 March.

23. Elek, A., Yildiz, H.S., Akca, B., Oren, N.C., & Gundogdu, B. (2025). Evaluating the Efficacy of Perplexity Scores in Distinguishing AI-Generated and Human-Written Abstracts. *Academic Radiology*, 32(4), 1785–1790. <https://doi.org/10.1016/j.acra.2025.01.017> EDN: HMLYED
24. Jiang, S., & Li, K. (2025). Another Kind of Authenticity: the Visual Simulacra of Artificial Intelligence. *Digital Creativity*, 36(1), 75–91. <https://doi.org/10.1080/14626268.2025.2467087> EDN: ZRIUEB
25. Konzack, L. (2025). Generative AI, Simulacra, and the Transformation of Media Production. *Athens Journal of Mass Media & Communications*, 11(3). <https://doi.org/10.30958/ajmmc.11-3-3> EDN: BYIUCE
26. Lyon, C.F. (2023). Fake Cases, Real Consequences: Misuse of ChatGPT Leads to Sanctions. *NY Litigator*, 28(2), 8–12.
27. Mahowald, K., Ivanova, A.A., et al. (2024). Dissociating Language and thought in Large Language Models: a Cognitive Perspective. *Trends in Cognitive Sciences*. <https://doi.org/10.1016/j.tics.2024.01.011> EDN: SKTATZ
28. Miralles-González, P., Huertas-Tato, J., Martín, A., & Camacho, D. (2025). Not All Tokens are Created Equal: Perplexity Attention Weighted Networks for AI Generated Text Detection. <https://doi.org/arXivpreprint arXiv:2501.03940>
29. Pudasaini, S., Miralles-Pechuán, L., Lillis, D., & Llorens Salvador, M. (2024). Survey on AI-generated Plagiarism Detection: The Impact of Large Language Models on Academic Integrity. *Journal of Academic Ethics*. <https://doi.org/10.48550/arXiv.2407.13105>
30. Ram, A. (2023). *Generative AI and Text Creation*. Preprint. <https://doi.org/10.48550/arXiv.2407.13105>

Сведения об авторах:

Новикова Марина Львовна, доктор филологических наук, профессор кафедры русского языка и лингвокультурологии Института русского языка, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *сфера научных интересов*: искусственный интеллект; машинный перевод; нейросеть; большая языковая модель; цифровое присутствие; *e-mail*: novikova-ml@rudn.ru
ORCID: 0000-0002-4673-067X, SPIN-код: 1502-9636, AuthorID: 499765.

Новиков Филипп Николаевич, кандидат филологических наук, доцент кафедры иностранных языков юридического института, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, д. 6); *сфера научных интересов*: искусственный интеллект; машинный перевод; нейросеть; большая языковая модель; цифровое присутствие; *e-mail*: novikov-fn@rudn.ru
ORCID: 0000-0003-4884-3659, Scopus: 57191418034, SPIN-код: 3523-2539, AuthorID: 671112.

Information about the authors:

Marina L. Novikova, Dr.Sc. (Philology), Professor, Russian Language and Cultural Studies Department, Russian Language Institute, RUDN University (6 Miklukho-Maklaya str., Moscow, Russian Federation 117198); *Research interests*: artificial intelligence; machine translation; neural networks; large language models; digital presence; *e-mail*: novikova-ml@rudn.ru
ORCID: 0000-0002-4673-067X, SPIN-code: 1502-9636, AuthorID: 499765.

Philipp N. Novikov, PhD in Philology, Associate Professor, Law Institute Department of Foreign Languages, RUDN University (6 Miklukho-Maklaya str., Moscow, Russian Federation 117198); *Research interests*: artificial intelligence; machine translation; neural networks; large language models; digital presence; *e-mail*: novikov-fn@rudn.ru
ORCID: 0000-0003-4884-3659, Scopus: 57191418034, SPIN-code: 3523-2539, AuthorID: 671112.