



оригинальная статья

<https://elibrary.ru/enimqp>

Искусственная нейросетевая генерация текста как новый вид маскировки письменной речи автора

Огорелков Игорь Витальевич

Государственный институт русского языка им. А. С. Пушкина, Россия, Москва

eLibrary Author SPIN: 5323-8641

<https://orcid.org/0000-0001-5116-3975>

ogorelkov.69@mail.ru

Махаева Елизавета Владиславовна

Московский Исследовательский Центр, Россия, Москва

<https://orcid.org/0009-0006-5098-621X>

Аннотация: Цель статьи – разработать новый вид маскировки письменной речи автора, такой как *искусственная нейросетевая генерация текста*, для проведения автороведческих экспертиз и последующих исследований. Материалом послужили тексты, выполненные человеком, и тексты, сгенерированные большой языковой моделью GPT 4.5. Применены методы композиционно-семантического, структурно-семантического и грамматико-синтаксического анализа текстов, порожденных нейросетью в результате выполнения задачи по маскировке письменной речи автора. Данные методы позволили выявить и описать дефекты сгенерированных текстов, обусловленные нарушениями в их имплицативной и референциальной семантике. Раскрыты актуальные вопросы языкового моделирования, возможностей генеративных языковых моделей, основанных на нейросетевых алгоритмах. Отмечена необходимость разграничения текстов естественного происхождения и генеративной природы для установления авторства. Установлено наличие корреляции между наполнением промпта дополнительными параметрами и результатом решения языковой моделью задачи по выполнению маскировки письменной речи автора. Сформулирован перечень типовых языковых признаков сгенерированных текстов. Приведены примеры сравнения текстов-оригиналов и текстов, порожденных нейросетью. Перспективы дальнейших исследований заключаются в углубленном изучении возможностей построения нейросетью сложных текстовых структур, включающих черты идиостиля автора текста-оригинала, предварительно загруженного в модель, по заданным параметрам промптов, а также формирование диагностического комплекса признаков искусственной нейросетевой генерации текста как нового вида маскировки, в основу которого будет положен перечень типовых языковых признаков искусственно сгенерированных текстов.

Ключевые слова: автороведческая экспертиза, нейросети, генерация текста, письменная речь, маскировка письменной речи, языковые модели

Цитирование: Огорелков И. В., Махаева Е. В. Искусственная нейросетевая генерация текста как новый вид маскировки письменной речи автора. *Виртуальная коммуникация и социальные сети*. 2025. Т. 4. № 2. С. 163–171. <https://doi.org/10.21603/2782-4799-2025-4-2-163-171>

Поступила в редакцию 12.03.2025. Принята после рецензирования 05.05.2025. Принята в печать 06.05.2025.

full article

Artificial Neural Network Text Generation as a New Type of Masking the Authorship

Igor V. Ogorelkov

Pushkin State Russian Language Institute, Russia, Moscow

eLibrary Author SPIN: 5323-8641

<https://orcid.org/0000-0001-5116-3975>

ogorelkov.69@mail.ru

Elizaveta V. Makhaeva

Moscow Research Center, Russia, Moscow

<https://orcid.org/0009-0006-5098-621X>

Abstract: Artificial neural network text generation may be used to conceal the authorship of a text. As a result, experts have to be able to tell the difference between natural and generative texts to establish authorship. In this research, human-made texts and texts generated by GPT 4.5 were analyzed using the methods of compositional-semantic, structural-semantic, and grammatical-syntactic analysis. The analysis made it possible to identify and describe some typical flaws in the AI-generated texts caused by violations in their implicative and referential semantics. It revealed some current issues of language modeling and the possibilities of generative language models based on neural network algorithms. The quality of the generated text depended on additional parameters to the prompt. The article describes linguistic features typical of generated texts and provides each case with comparative examples. The prospects for further research lie in an in-depth study of complex AI-generated text structures, e.g., preloaded idiosyncratic texts. The practical result may yield a diagnostic complex of signs based on a list of typical linguistic features of AI-generated texts.

Keywords: authorship expertise, neural networks, text generation, written speech, masking of written speech, language models

Citation: Ogorelkov I. V., Makhaeva E. V. Artificial Neural Network Text Generation as a New Type of Masking the Authorship. *Virtual Communication and Social Networks*, 2025, 4(2): 163–171. (In Russ.) <https://doi.org/10.21603/2782-4799-2025-4-2-163-171>

Received 12 Mar 2025. Accepted after review 5 May 2025. Accepted for publication 6 May 2025.

Введение

В современном мире искусственный интеллект (ИИ) является данностью, функционально доступной практически каждому рядовому пользователю сети Интернет. Нейронные сети как неотъемлемый и активно модернизируемый инструмент ИИ получил широкое распространение из-за простоты и доступности использования.

Искусственные нейронные сети представляют собой одно из научных направлений исследований в области ИИ, которое основано на попытках воспроизвести нервную систему человека, а именно способность программы обучаться и исправлять ошибки [Zupan, Gasteiger 1991: 14]. Тем самым нейросети, подобно системе нейронов головного мозга, представлены структурой нейронов иного характера – вычислительных элементов, каждый из которых имеет несколько входов-синапсов

и один выход-аксон [Бурнашев, Аламова 2023: 259]. В настоящее время нейросети применяются для решения широкого ряда задач, в том числе лингвистических: анализ настроений (*sentiment analysis*), распознавание речи (*speech recognition*), информационный поиск (*data searching*), машинный перевод (*machine translation*) и др. Данные задачи традиционно объединяются исследователями в междисциплинарное направление, вобравшее в себя принципы машинного обучения и лингвистики, – *Natural Language Processing* (NLP) [Тельпов, Ларцина 2023: 47]. Одна из областей применения NLP – генерация текстов разнообразных жанров, объемов и свойств посредством использования таких генеративных языковых моделей, как большие языковые модели (*Large Language Model* – LLM).

Стремительное развитие LLM стало возможным благодаря достижениям в области машинного обучения, увеличению вычислительных мощностей и совершенствованию алгоритмов оптимизации [Zhang et al. 2021: 108; Vaswani et al. 2017]. Изначально обработка естественного языка базировалась на статистических методах – модели скрытых марковских процессов и n-граммных моделях [Jurafsky, Martin 2009: 211]. Однако такие подходы были ограничены в понимании контекста и генерации осмысленных текстов. Прорыв в этой области произошел с появлением нейронных сетей и особенно рекуррентных нейронных сетей (*Recurrent Neural Network* – RNN), которые позволили моделям учитывать последовательную природу текстовых данных [Sutskever et al. 2014: 3107].

Наиболее распространенные и удобные для пользователей генеративные языковые модели, разработанные компаниями OpenAI и Google, – GPT и BERT – основаны на нейросетевых алгоритмах. Указанные модели уже продемонстрировали способность не только к воспроизведению естественного языка, но и его генерации на уровне, близком к человеческому. По мере продолжающегося роста технологических возможностей генеративных языковых моделей актуализируется и проблема переосмысления статуса человека как единоличного, безальтернативного субъекта создания текстов на естественном языке. Это влечет за собой зарождение новых постгуманистических, постчеловеческих концепций существования и употребления языка и дискурсивных практик в сфере теоретических лингвистических исследований [Зырянова, Чернавский 2024: 145]. Вместе с тем российские исследователи отмечают роль генеративных языковых моделей в медиакультуре XXI в. и их влияние на процессы коммуникации [Солдаткина, Чернавский 2023: 44], а также их использование в сфере цифровизации нормотворчества [Гуляева 2023: 115].

В то же время повышение доступности, увеличение степени распространения и, как следствие, бесконтрольный характер использования больших языковых моделей провоцирует соответствующие вызовы для проведения прикладных исследований. Одним из таких вызовов является область автороведческой экспертизы, решающей в том числе диагностическую задачу выявления признаков маскировки письменной речи автора.

В автороведении под маскировкой письменной речи понимается сознательное изменение автором

текста формально-содержательной (смысловой) и языковой структуры своей письменной речи с целью скрыть от адресата информацию о своей личности или дать заведомо ложную информацию [Рубцова и др. 2013: 23]. На сегодняшний день маскировка письменной речи условно разделяется на три основных вида: подстройка под письменную речь конкретного лица, искажение письменной речи (за счет снижения уровня грамотности, изменения уровня культуры речи, нейтрализации письменной речи), имитация авторства (женщины, мужчины, лица молодого, пожилого возраста и конкретных профессий) [Огорелков 2020: 67].

Однако ситуация повсеместного использования нейросетевых чат-ботов, которые решают задачу генерации текстов, порождает возможность их применения, например, в деструктивных целях сокрытия личности автора и создания конфликтогенных текстов от имени иного лица. При этом на данном этапе развития лингвистика и автороведение не располагают разработанным инструментарием (устойчивый комплекс признаков), позволяющим разграничить естественные и сгенерированные тексты, а также апробированным на достаточном корпусе текстов данных видов и доступных для применения специалистов без использования дополнительных технических средств.

Таким образом, цель статьи – разработать новый вид маскировки письменной речи автора, такой как *искусственная нейросетевая генерация текста*. Под ней мы понимаем составление текста документа от имени другого лица с использованием текстового генератора. Для достижения поставленной цели необходимо, во-первых, выявить возможности генеративных языковых моделей к выполнению маскировки предварительно загруженной в них письменной речи автора текста. Во-вторых, установить типовые языковые признаки искусственно сгенерированных текстов, которые при проведении дальнейших исследований будут положены в основу диагностического комплекса признаков искусственной нейросетевой генерации текста как нового вида маскировки.

Несмотря на то что специалисты в области создания и применения LLM, постоянно совершенствуя данный инструмент, уделяют много внимания настройке параметров генеративной языковой модели [Фищева и др. 2023], разработке эффективных методов управления генерацией текста [Ostyakova et al. 2023], в том числе и контролю качества коротких текстов [Кривошеев и др. 2021: 118],

оценке уровня доверия и др., сгенерированные нейросетями тексты все же характеризуются наличием определенных семантических дефектов, называемых специалистами в области NLP *галлюцинациями*. Под ними понимаются определенные нарушения в семантическом устройстве сгенерированных текстов, т.е. несоответствие между входными данными и полученным результатом, фактологические отклонения, нерелевантные фрагменты в сгенерированных текстах [Maynez et al. 2020: 1907; Floridi, Chiriatti 2020: 685]. Кроме того, учеными отмечается неспособность языковых моделей выстраивать долгосрочные семантические зависимости в тексте, что иногда приводит к «вырождению» текста [Mikhaylovskiy, Churilov 2023].

Тем не менее большие языковые модели, основанные на нейросетевых алгоритмах, вполне успешно демонстрируют возможности генерации связного и когерентного текста, оперативно реагируя на запрос пользователя и существенно меняя дискурсивные, жанровые и смысловые характеристики текста при внесении минимальных корректировок в предмет запроса. Проблемы порождения когерентного текста генеративными моделями детально исследовались А. В. Козловским, Я. Э. Мельником, В. И. Волощуком. Они отмечают относительно постоянное выявление изъянов сгенерированных текстов при наиболее подробном анализе, а также указывают на существенный недостаток автоматической генерации текстов – учет исключительно семантики слов, что закономерно создает проблемы с логичностью повествования и, следовательно, требует учета контекстуального окружения спорных лексем и словосочетаний [Козловский и др. 2022: 161].

Непрекращающееся технологическое развитие нейросетевых алгоритмов, совершенствование параметров генерации текста, способность языковых моделей к самообучению позволяют значительно сократить сроки и увеличить качество процесса по устранению семантических дефектов в сгенерированных текстах. Однако работа с практическим материалом в ходе проводимого нами исследования показала, что сгенерированные тексты все же характеризуются неточностью фактологической информации, нарушением лексической валентности слов, семантической аномальностью и иными маркерами, позволяющими отличить текст генеративной природы от текста естественного происхождения.

Методы и материалы

Материал исследования – письменные тексты, разделенные нами на два кластера: первый – тексты бытового и политического дискурса, выполненные человеком, выявленные в том числе с использованием Национального корпуса русского языка ($n = 100$); второй – тексты, сгенерированные большой языковой моделью GPT 4.5, основанной на нейросетевых алгоритмах ($n = 100$).

Модель GPT была разработана OpenAI в 2018 г. Благодаря постоянному усовершенствованию на сегодняшний день она является наиболее мощным инструментом для обработки и генерации текстов на естественном языке (последняя версия – GPT 4.5). Данная нейросеть состоит из трансформаторов нескольких уровней, которые обрабатывают естественные тексты и генерируют на их основе когерентные выходные данные [Тельпов, Ларцина 2023: 48], и обучается в два этапа: предобучение на большом объеме текста и дообучение на конкретных задачах [Radford et al. 2018].

Технология проводимого нами исследования включала два этапа:

I. Выполнение маскировки письменной речи автора посредством формулирования запроса для языковой модели GPT 4.5. Используемые тексты естественного происхождения (порожденные человеком) цельные, связные, завершенные; выполнены на русском языке; вербальный объем включает не менее 100 словоформ.

II. Сравнение естественного и сгенерированного текстов, в ходе которого выявлены возможности генеративных языковых моделей к выполнению маскировки письменной речи автора текста, предварительного загруженного в модель, и установлено наличие типовых языковых признаков искусственно сгенерированных текстов.

Применены методы композиционно-семантического, структурно-семантического и грамматико-синтаксического анализа текстов, сгенерированных в результате выполнения задачи по речевой маскировке. Данные методы позволили выявить и описать дефекты сгенерированных текстов, порожденные нарушениями в их имплицативной и референциальной семантике.

Результаты

В рамках исследования специального подхода потребовалось формулирование промпта для большой языковой модели GPT 4.5, содержащего указание произвести маскировку письменной речи текста-

оригинала, загружаемого в модель. Промпт – запрос, инструкция-подсказка, посредством которой ставится задача для языковой модели [Глазова и др. 2024]. Так, с помощью естественного языка пользователем формулируется, задается роль и отправляется на вход модели текстовый запрос, который она обрабатывает, а затем подбирает наиболее вероятный и точный ответ, соответствующий входным данным [Стецик 2024: 157].

Таким образом, именно от качества сформулированного промпта зависит результат генерации текста моделью. Исследователи отмечают, с одной стороны, необходимость формулирования распространенного промпта путем включения в его состав коммуникативной роли, которую модели необходимо реализовать, указания контекста коммуникации, приведения примеров для облегчения процесса генерации [Стецик 2024: 159], а с другой – указывают на важность лаконичного и структурированного высказывания, закладываемого в основу промпта [Алексеева, Алексеев 2024: 53].

Промпты были условно разделены:

1. Простые: ограничивались побудительным высказыванием, содержащим установку на генерацию нового текста от имени автора текста-оригинала, загруженного в модель, и имели следующее содержание: *Сгенерируй текст от имени автора текста, содержащегося в файле «оригинал.docx».*

2. Сложные: формулировались с обязательным изменением темы текста-оригинала, указанием роли, установкой на сохранение орфографических, пунктуационных, лексических (особенно в случае конфликтогенной направленности текста-оригинала) особенностей идиостиля автора первоначального текста, и имели следующее содержание: *Представь, что ты автор текста, содержащегося в файле «оригинал.docx». Напиши жалобу на мальчика, который обижает сына автора текста, имитируя его идиостиль. Сохраняй орфографические и пунктуационные ошибки автора изначального текста. Не будь вежлив.*

Проведение подобной дифференциации промптов на основании наличия / отсутствия задаваемых в них параметров позволило подтвердить прямую зависимость между наполнением промпта и результатом выполнения языковой моделью маскировки письменной речи автора первоначального текста посредством генерации нового. Так, при использовании простого промпта модель генерировала семантически аномальный текст (табл. 1).

Данные таблицы 1 показывают, что в результате выполнения моделью маскировки письменной речи автора порожденный нейросетью текст, напротив, утратил черты авторского идиостиля, хотя и остался правильным в структурно-семантическом отношении. Общая структура сгенерированного текста, а также имеющаяся в нем избыточность анцеляризов и штампов свойственна официально-деловому или публицистическому функциональному стилям. Текст-оригинал же выполнен в разговорно-обиходном стиле. Данное явление можно объяснить тем, что, во-первых, модель обучалась на медицинских, юридических и экономических текстах; во-вторых, промпт не содержал конкретных указаний относительно сохранения черт идиостиля автора первоначального текста, его орфографических, пунктуационных, лексико-грамматических особенностей.

В целом модель способна сгенерировать грамматически и композиционно-семантически правильный, т.е. когерентный текст, построенный по модели текста-оригинала. При этом порожденной нейросетью текст содержит значительное количество семантических аномалий (*Ваши действия явно продиктованы злом; Я как равнодушный гражданин и человек, чувствующий себя рабом Божиим*) и грамматически аномальных предложений (*Почему остановка у жилого дома оказалась менее важной?; Явно, что вы отмыли деньги*).

В свою очередь результат решения языковой моделью задачи по маскировке письменной речи автора при применении сложного промпта был более релевантен. Так, сгенерированный текст характеризуется не только структурно-семантической правильностью, когерентностью, отсутствием семантических аномалий, но и сохраняет черты идиостиля автора текста-оригинала, о чем свидетельствует наличие жаргонизмов и просторечий на лексическом языковом уровне (*шараге, столовке*), пунктуационных особенностей (отсутствие знаков препинания в конце предложений, входящих в состав числового списка), элементов разговорно-обиходного функционального стиля (табл. 2).

В результате сравнения текстов естественного происхождения и текстов, сгенерированных языковой моделью (посредством использования как простых, так и сложных промптов) в ходе решения задачи по маскировке письменной речи автора, нам удалось сформулировать перечень типовых языковых признаков текстов, порожденных нейросетью:

Табл. 1. Фрагмент генерации текста посредством применения простого промпта
Tab. 1. Text generated using a simple prompt

Оригинал	Нейросеть
<p>В Департамент транспорта и развития дорожно-транспортной инфраструктуры¹</p> <p>Вы, продажные чиновники, обещали нам новые автобусы. А в итоге что? Их стало меньше. Если раньше у нас было четыре автобуса (740, 280, 742, 738), то сейчас их три, а по выходным вообще два. Вы серьезно думаете, что если вы смените номера автобусов все сразу плясать начнут и никто ничего не заметит? И сколько же денег вы, безбожники, отмыли, чтобы просто изменить номера автобусов? Теперь из-за того, что вы убрали 742 автобус мне приходится добираться на двух автобусах до электрички с сумками. А это не так уж и просто. И знаете что? Мне нравится, что вы решили не останавливаться на автобусах. Вы решили мне окончательно испортить жизнь и убрали остановку около моего дома. Скажите мне пожалуйста, чем вы думали. То есть по вашей логике остановка, которая находится у гаражей, где никогда никого нет куда нужнее, чем остановка у жилого дома? Вами руководит Сатана!</p> <p><...></p> <p>Неравнодушный гражданин и раб божий</p>	<p>Департамент транспорта и развития дорожной инфраструктуры,</p> <p>Я, как неравнодушный гражданин и человек, чувствующий себя рабом Божьим, вынужден выразить свое возмущение по поводу текущей ситуации с транспортом в нашем районе. Ваши обещания о новых автобусах оказались лишь красивыми словами, потому что на деле их стало меньше. Ранее у нас было четыре маршрута (740, 280, 742, 738), а теперь осталось только три, и то по выходным – всего два. Вы действительно считаете, что смена номеров автобусов изменит ситуацию и никто не заметит, что стало хуже? Явно, что вы отмыли деньги, заменяя номера автобусов вместо реального улучшения транспортного обслуживания. Теперь из-за исчезновения 742 автобуса мне приходится добираться до электрички на двух автобусах с сумками, что крайне неудобно. А еще вы решили лишить нас остановки у моего дома, заменив ее на остановку, которая находится у гаражей, где никогда никого нет. Как вы можете объяснить такую нелепость? Почему остановка у жилого дома оказалась менее важной? Ваши действия явно продиктованы злом.</p> <p><...></p> <p>Неравнодушный гражданин и раб божий</p>

Табл. 2. Фрагмент генерации текста посредством применения сложного промпта
Tab. 2. Text generated using a complex prompt

Оригинал	Нейросеть
<p>Кротову Д. С., ректору вуза</p> <p>В вашей шараге происходит полный беспредел!</p> <p>1. сначала вы набираете много студентов, а после сбора денег отчисляете их с первого или второго курса</p> <p>2. бюджетные места достаются «элите» – детям судей</p> <p>3. преподаватели:</p> <ul style="list-style-type: none">– старики, которые ведут информатику, а сами не умеют пользоваться компьютером– не дают знаний, не читают лекции в полном объеме, а на семинарах спрашивают, как с профессоров <p><...></p> <p>Кротов, жди проверок!</p> <p>Такое же письмо я отправил в Министерство образования и науки РФ, Департамент образования города Москвы, чтобы закрыли наконец эту лавочку по сбору денег!</p>	<p>Кротову Д. С., ректору вуза</p> <p>В вашей шараге происходит полный беспредел!</p> <p>1. еда в столовой отвратительная – блюда несвежие, иногда даже тухлые</p> <p>2. меню однообразное, студентам приходится давиться одним и тем же каждый день</p> <p>3. цены неоправданно завышены, за такие бабки можно поесть в нормальном кафе</p> <p>4. персонал грубит студентам, не соблюдает санитарные нормы</p> <p><...></p> <p>Кротов, жди проверок!</p> <p>Такое же письмо я отправил в Роспотребнадзор и Департамент здравоохранения города Москвы, чтобы навели наконец порядок в вашей столовке!</p>

¹ Здесь и далее авторские орфография и пунктуация сохранены.

1. Избыточность штампов и канцеляризмов (*В связи с вышеописанным; Позвольте мне выразить глубокую озабоченность; Дело в том, что; В целях доведения до вашего сведения; Невозможно не задаться вопросом*); лишних словоформ вне требований контекста (*занимается скрытым наблюдением вместо тайне наблюдает*).

2. Нарушение правил использования степени сравнения прилагательного (*Остановка у жилого дома оказалась менее важной*); лексической валентности слов (*Ваши действия явно продиктованы злом*); грамматической сочетаемости при использовании разговорных слов и жаргонизмов (*Еще раз ты припаркуешь свою помойку здесь, и одной запиской под дворниками это не отделаешься, мудила*); правил употребления предлогов перед согласными (*С второго раза это начинает дико бесить*).

3. Наличие случаев языковой интерференции – дословного перевода синтаксической конструкции с английского языка на русский (*Наш разговор с ним потерял контроль как Our communication with him lost control*); значительного количества вводных слов и конструкций (*Кроме того; Вместе с тем; Вероятнее всего; По всей видимости; Представляется*).

Следует также отметить, что ряд выявленных признаков не связан с проявлением семантической аномальности, а обусловлен объективными причинами функционирования языковой модели GPT (ее базовым языком является английский). В связи с этим в числе выявленных признаков наблюдаются случаи межъязыковой интерференции, нарушение правил употребления предлогов перед согласными и т. п.

Заключение

В ходе проведенного экспериментального исследования было установлено наличие корреляции между наполнением промпта дополнительными конкретизирующими параметрами и результатом решения языковой моделью задачи по выполнению маскировки письменной речи автора первоначального текста посредством генерации нового. Так, использование простого промпта при постановке перед языковой моделью задачи по выполнению речевой маскировки показало свою несостоятельность: сгенерированный текст характеризуется наличием семантических дефектов и аномалий и практически не содержит

характерных черт идиостиля автора текста-оригинала. При этом формулирование сложного промпта позволяет достичь качественно релевантного результата генерации: модель успешно выполняет речевую маскировку, но текст может содержать избыточность штампов и канцеляризмов, элементы языкового калькирования и т. д.

Вместе с тем стоит отметить, что в настоящий момент большая языковая модель GPT 4.5 демонстрирует способность к генерации текстов, отличающихся грамматической, структурно-семантической и композиционной правильностью, при этом независимо от сложности заданного промпта. Данное обстоятельство свидетельствует о наличии потенциальной возможности выполнения маскировки письменной речи автора в целях сокрытия своей личности посредством применения нейросетевых генеративных инструментов рядовым пользователем сети Интернет. Это актуализирует проведение дальнейшей разработки искусственной нейросетевой генерации текста как нового вида маскировки письменной речи автора.

Перспективы дальнейших исследований заключаются в углубленном изучении возможностей построения нейросетью сложных текстовых структур, включающих черты идиостиля автора текста-оригинала, предварительно загруженного в модель, по заданным параметрам промптов, а также формирование диагностического комплекса признаков искусственной нейросетевой генерации текста как нового вида маскировки, в основу которой будет положен перечень типовых языковых признаков искусственно сгенерированных текстов.

Конфликт интересов: Авторы заявили об отсутствии потенциальных конфликтов интересов в отношении исследования, авторства и / или публикации данной статьи.

Conflict of interests: The authors declared no potential conflict of interests regarding the research, authorship, and / or publication of this article.

Критерии авторства: Авторы в равной степени участвовали в подготовке и написании статьи.

Contribution: All the authors contributed equally to the study and bear equal responsibility for the information published in this article.

Литература / References

- Алексеева Л. Г., Алексеев П. С. Язык промптов, или особенности формулирования запросов к генеративным нейросетям для создания изображений. *Verba. Северо-Западный лингвистический журнал*. 2024. № 3. С. 50–61. [Alekseeva L. G., Alekseev P. S. Prompt language, or features of formulation of queries to generative neural networks for image creation. *Verba*, 2024, (3): 50–61. (In Russ.)] [https://doi.org/10.34680/VERBA-2024-3\(13\)-50-61](https://doi.org/10.34680/VERBA-2024-3(13)-50-61)
- Бурнашев Р. Ф., Аламова А. С. Роль нейронных сетей в лингвистических исследованиях. *Science and Education*. 2023. Т. 4. № 3. С. 258–269. [Burnashev R. F., Alamova A. S. The role of neural networks in linguistic research. *Science and Education*, 2023, 4(3): 258–269. (In Russ.)]
- Глазова Л. И., Лузгина А. Д., Пугачевский А., Кочетова А. Н., Фейзуллов Д., Чиж А. В., Виноградов М. Ю. Искусственный интеллект как эффективный инструмент коммуникаций. *Российская школа связей с общественностью*. 2024. № 33. С. 48–65. [Glazova L. I., Luzgina A. D., Pugachevsky A., Kochetova A. N., Feyzullov D., Chizh A. V., Vinogradov M. Yu. Artificial intelligence as an effective communication tool. *Rossiiskaia shkola svyazei s obshchestvennostiu*, 2024, (33): 48–65. (In Russ.)] <https://elibrary.ru/jurxuo>
- Гуляева П. С. Применение генеративных языковых моделей в условиях цифровизации нормотворчества. *Вестник МГПУ. Серия: Юридические науки*. 2023. № 3. С. 126–137. [Gulyaeva P. S. Large language models use in the context of rulemaking digitalization. *The Academic Journal of Moscow City University. Series "Legal Sciences"*, 2023, (3): 126–137. (In Russ.)] <https://doi.org/10.25688/2076-9113.2023.51.3.11>
- Зырянова И. Н., Чернавский А. С. Генеративные языковые модели и феномен антиантропоцентризма – новые перспективы лингвистической парадигмы «posthumano» и «общего/сильного» ИИ. *Известия Байкальского государственного университета*. 2024. Т. 34. № 1. С. 144–152. [Ziryanova I. N., Chernavskiy A. S. Generative language patterns and the phenomenon of anti-anthropocentrism – new perspectives on the linguistic paradigm of "posthumano" and "general/strong" AI. *Bulletin of Baikal State University*, 2024, 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)
- Козловский А. В., Мельник Я. Э., Волощук В. И. О подходе для автоматической генерации сюжетно связанного текста. *Известия Тульского государственного университета. Технические науки*. 2022. № 9. С. 160–167. [Kozlovsky A. V., Melnik Ya. E., Voloshchuk V. I. On the approach for automatic generation of narrative-linked text. *Izvestiya Tula State University*, 2022, (9): 160–167. (In Russ.)] <https://elibrary.ru/dirfga>
- Кривошеев Н. А., Иванова Ю. А., Спицын В. Г. Автоматическая генерация коротких текстов на основе применения нейронных сетей LSTM и SeqGAN. *Вестник Томского государственного университета. Управление, вычислительная техника и информатика*. 2021. № 57. С. 118–130. [Krivoshcheev N. A., Ivanova Yu. A., Spitsyn V. G. Automatic generation of short texts based on the use of neural networks LSTM and SeqGAN. *Tomsk State University Journal of Control and Computer Science*, 2021, (57): 118–130. (In Russ.)] <https://doi.org/10.17223/19988605/57/13>
- Огорелков И. В. Особенности диагностического автороведческого исследования анонимного документа на основании признаков, характеризующих имитацию авторства. *Русский язык за рубежом*. 2020. № 1. С. 66–69. [Ogorelkov I. V. Peculiarities of the diagnostic author's study of anonymous document on the basis of signs, characterizing the authority imitation. *Russkii iazyk za rubezhom*, 2020, (1): 66–69. (In Russ.)] <https://elibrary.ru/seg1hn>
- Рубцова И. И., Ермолова Е. И., Безрукова А. И., Огорелков И. В. Установление факта маскировки письменной речи в тексте анонимного документа: методические рекомендации. М.: ЭКЦ МВД России, 2013. 64 с. [Rubtsova I. I., Ermolova E. I., Bezrukova A. I., Ogorelkov I. V. *Establishing the fact of masking written speech in the text of an anonymous document: Methodological recommendations*. Moscow: EKTS MVD Rossii, 2013, 64. (In Russ.)]
- Солдаткина Я. В., Чернавский А. С. Генеративные языковые модели как актуальный феномен медиакультуры в начале XXI века. *Наука и школа*. 2023. № 4. С. 42–54. [Soldatkina Ya. V., Chernavskiy A. S. Generative language models as a crucial phenomenon of media culture at the beginning of the XXI century. *Science and School*, 2023, (4): 42–54. (In Russ.)] <https://doi.org/10.31862/1819-463X-2023-4-44-56>

- Стецки М. Союз лингвистики и промт-инжиниринга: лингвистические особенности запросов к нейросети. *Vilnius University Open Series*. 2024. С. 155–166. [Stetsyk M. The union of linguistics and prompt engineering: Linguistic features of prompts to neural networks. *Vilnius University Open Series*, 2024: 155–166. (In Russ.)] <https://doi.org/10.15388/SV-I-II.2024.14>
- Тельпов 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. [Telpov R. E., Lartsina S. V. Typological differences of natural and neural network-generated texts in a quantitative aspect. *Nauchnyi dialog*, 2023, 12(7): 258–269. (In Russ.)] <https://doi.org/10.24224/2227-1295-2023-12-7-47-65>
- Фищева И. Н., Пескишева Т. А., Головизнина В. С., Котельников Е. В. Метод классификации аспектов аргументации в русскоязычных текстах. *Программные системы: теория и приложения*. 2023. Т. 14. № 4. С. 25–45. [Fishcheva I. N., Peskisheva T. A., Goloviznina V. S., Kotelnikov E. V. A method for classifying aspects of argumentation in Russian-language texts. *Program systems: Theory and applications*, 2023, 14(4): 23–45. (In Russ.)] <https://doi.org/10.25209/2079-3316-2023-14-4-25-45>
- Floridi L., Chiriatti M. GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 2020, 30(4): 681–694. <https://doi.org/10.1007/s11023-020-09548-1>
- Jurafsky D., Martin J. H. *Speech and Language Processing*. Pearson Prentice Hall, 2009, 988.
- Maynez J., Narayan S., Bohnet B., McDonald R. On faithfulness and factuality in abstractive summarization. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2020, 1906–1919. <https://doi.org/10.18653/v1/2020.acl-main.173>
- Mikhaylovskiy N., Churilov I. Autocorrelations decay in texts and applicability limits of language models. *Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference "Dialogue"*, 2023, (22), 350–360. <https://doi.org/10.48550/arXiv.2305.06615>
- Ostyakova L., Petukhova K., Smilga V., Zharikova D. Linguistic annotation generation with ChatGPT: A synthetic dataset of speech functions for discourse annotation of casual conversations. *Computational Linguistics and Intellectual Technologies: Papers from the Annual International Conference "Dialogue"*, 2023, (22), 386–403.
- Radford A., Narasimhan K., Salimans T., Sutskever I. Improving language understanding by generative pre-training. *OpenAI*, 2018. URL: https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf (accessed 10 Mar 2025).
- Sutskever I., Vinyals O., Le Q. V. Sequence to sequence learning with neural networks. *Advances in Neural Information Processing Systems 27*, 2014, 3104–3112.
- Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A. N., Kaiser L., Polosukhin I. Attention is all you need. *Advances in Neural Information Processing Systems 30*, 2017, 5998–6008.
- Zhang C., Bengio S., Hardt M., Recht B., Vinyals O. Understanding deep learning (still) requires rethinking generalization. *Communications of the ACM*, 2021, 64(3): 107–115. <https://doi.org/10.1145/3446776>
- Zupan J., Gasteiger J. Neural networks: A new method for solving chemical problems or just a passing phase? *Analytica Chimica Acta*, 1991, 248(1): 1–30. [https://doi.org/10.1016/S0003-2670\(00\)80865-X](https://doi.org/10.1016/S0003-2670(00)80865-X)