

Научная статья

УДК 81.374

DOI: <https://doi.org/10.18721/JHSS.14405>



ГРТ И CHATGPT: КАК ЭТО ПО-РУССКИ? (ОБ ЭЛЕКТРОННЫХ СЛОВАРНЫХ СТАТЬЯХ В НЕОЛОГИЧЕСКОМ РЕСУРСЕ NEOLEX.ILING.SPB.RU)

Н.В. Козловская 

Институт лингвистических исследований РАН,
Санкт-Петербург, Российская Федерация

 mnegolosbyl@gmail.com

Аннотация. В статье проанализирована проблема лексикографирования новых заимствований, происходящих от английского сращения «ChatGPT», в оперативном цифровом неологическом ресурсе «Новое в русской лексике. Словарные материалы» (neolex.iling.spb.ru). В процессе анализа словоупотреблений и составления конкорданса обнаружено, что в качественной прессе используется оригинальное имя собственное в исходной графической форме «Chat GPT». За пределами качественной прессы – в письменной коммуникации специалистов в сфере бизнеса, стартапов, инноваций, маркетинга и новых технологий – выявлены многочисленные варианты неологических единиц («чатгпт», «чатГПТ», «чат ГПТ», «ГРТ-чат», «джипити», «жипити», «ДжиПиТи») и их производные («джипитишный», «джипитичный», «джипитить», «джипитишка»). Охарактеризованы разные типы вариативности, отраженные в конкордансе: фонетико-графическая, орфографическая, грамматическая, выявлена связанность морфологической вариантности, семантического обогащения и варьирования. Проанализированы трудности, возникающие в процессе лексикографирования неологизмов: наличие компонентов, терминологическая природа, пограничный характер, связанный с переходом из имен собственных в аппеллятивы, отсутствие кодифицированного варианты. Разработаны рекомендации по созданию электронной словарной статьи с включением всех выявленных вариантов. Показано, что отказ от лексикографирования всех частотных написаний вступает в противоречие с реальной языковой практикой. В статье отражена специфика уникального неологического ресурса neolex.iling.spb.ru, в котором могут отражаться слова, не прошедшие этапа кодификации. Динамический характер базы neolex.iling.spb.ru, в которой реализована возможность оперативно изменять данные словарной статьи, позволит после кодификации одного из вариантов написания характеризуемых неологических единиц внести коррективы в зоны вокабулы, орфографической и энциклопедической справок.

Ключевые слова: неография, цифровой лексикографический ресурс, электронная словарная статья, имя собственное, аппеллятив, вариант, заимствование.

Для цитирования: Козловская Н.В. ГРТ и CHATGPT: как это по-русски? (об электронных словарных статьях в неологическом ресурсе neolex.iling.spb.ru) // Terra Linguistica. 2023. Т. 14. № 4. С. 67–78. DOI: 10.18721/JHSS.14405



GPT AND CHATGPT: HOW IS IT IN RUSSIAN? (ON ELECTRONIC DICTIONARY ENTRIES IN THE NEOLOGICAL DATABASE NEOLEX.ILING.SPB.RU)

N.V. Kozlovskaja 

The Institute for Linguistic Studies of the Russian Academy of Sciences,
St. Petersburg, Russian Federation

 mnegolosbyl@gmail.com

Abstract. The article analyzes the lexicography of new borrowings originating from the English word “ChatGPT”. As borrowing has spread, it is necessary to timely reflect new Russian words in the digital neological resource “New in Russian vocabulary. Dictionary materials” (neolex.iling.spb.ru). During the concordance process, the author determined that so-called quality press uses the original English name “Chat GPT”. Numerous variants of neological units (“chatgpt”, “chatGPT”, “chat GPT”, “GPT-chat”, “dzhipiti”, “zhipiti”, “DzhiPiTi”) and their derivatives (“dzhipitishnyj”, “dzhipitichnyj”, “dzhipitit’”, “dzhipitishka”) were identified in the written communication of specialists in the field of business, marketing and new technologies. The article also characterizes different types of variability reflected in the concordance: phonetic-graphic, orthographic, grammatical; it reveals the connectedness of morphological variability, semantic enrichment and variation. The article shows the difficulties that arise in the process of lexicography of neologisms, such as: variation, words containing Latin and Cyrillic components, the terminological nature of neologisms, the transition from proper names to appellatives, the lack of codified variants. The author offered recommendations for creating an electronic dictionary entry including all the identified variants of the new word. The article shows that the non-inclusion of a number of spellings conflicts with real language practice. The article reflects the specifics of the unique neological resource neolex.iling.spb.ru, which can reflect words that have not passed the codification stage. The dynamic nature of the neolex.iling.spb.ru database, which implements the ability to quickly change the data of a dictionary entry, will allow, after codifying one of the spelling variants of the characterized neological units, adjustments to the areas of vocabulary, spelling and encyclopedic references.

Keywords: neography, digital lexicographic resource, electronic dictionary entry, proper name, appellative, variant, borrowing.

Citation: Kozlovskaja N.V., GPT and CHATGPT: How is it in Russian? (on electronic dictionary entries in the neological database neolex.iling.spb.ru), *Terra Linguistica*, 14 (4) (2023) 67–78. DOI: 10.18721/JHSS.14405

Введение

Динамика лексико-семантической системы языка напрямую зависит от событий, оказывающих значительное влияние на жизнь общества. К важнейшим инфоповодам, получающим широкое отражение в медиапространстве, относятся знаковые события в области разработки технологий искусственного интеллекта.

Так, на фоне резкого технологического скачка 2010-х, связанного с появлением искусственных нейронных сетей и разработкой методов глубокого машинного обучения, в периодике и в интернет-пространстве зафиксирован значительный рост частоты употребления словосочетания «искусственный интеллект» (в 2022 году рост замедляется в связи с появлением других инфоповодов и связанных с ним «неогенных зон»). Аналитические инструменты медиабанка «Интегрум» позволяют установить число словоупотреблений по данному запросу с учетом ежегодного роста [1].

Новым импульсом для языковой реакции становится распространение в медиапространстве цифровых исследований. О значимости процесса медиатизации IT-сферы пишет А.М. Плотни-

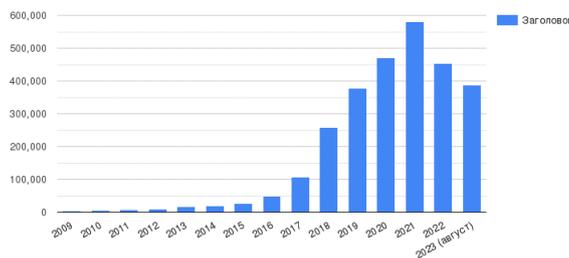


Рис. 1. Частота словоупотреблений сочетания «искусственный интеллект» с 2009 по 2023 гг. в медиабанке «Интегрум»

Fig. 1. Frequency of word usage for “artificial intelligence” from 2009 to 2023 in the Integrum media bank

кова: «Интерес к машинному обучению объясняется доступностью нейронных сетей, возможностью их использования в самых разных сферах человеческой жизни, следствием чего становится медиатизация терминов IT-сферы, обусловленная взаимодействием научного и медийного дискурсов» [2].

Важным событием 2023 года стало появление разработанной компанией OpenAI языковой модели GPT-4 с функцией обработки текста и картинок, а также широкое обсуждение в интернет-пространстве чат-бота с искусственным интеллектом. Популярность нейросети обусловила значительный рост использования английского имени собственного ChatGPT (Generative Pretrained Transformer), а также появление разнообразных вариантов обозначения в интернет-текстах разных жанров: чат ДжиПиТи, чат-джипити, GPT-чат, чатGPT, чатГПТ, чат ДЖПТ, чатЖПТ и некот. др.

Постановка проблемы

В первой четверти XXI в. современная лексикография переживает кардинальные преобразования под воздействием стремительно развивающихся цифровых технологий. Проблема специфики электронных словарей получила отражение в работах, посвященных современной цифровой лексикографии [3–9]. Называя интернет-словарь необходимым инструментом лингвиста, авторы статей пишут о кардинальных преобразованиях лексикографических принципов представления лингвистической информации и выявляют новые тенденции в формировании и функционировании различных инновационных лексикографических ресурсов.

К таковым относится разработанный в ИЛИ РАН цифровой лексикографический ресурс «Новое в русской лексике. Словарные материалы» (neolex.iling.spb.ru), реализующий относительно богатую функциональность в области поиска и структуризации словарной информации о новом слове. Данный ресурс, сочетающий инновационное и традиционное начала, опирается на принципы, разработанные при составлении выпусков «Новое в русской лексике» — т.н. словарей-ежегодников, выпуск которых после длительного перерыва был возобновлен в отделе лексикографии современного русского языка.

Основная задача ресурса заключается в оперативном отражении неологизмов; структура электронной словарной статьи наглядно отражает системный характер русской лексики и позволяет составителям быстро вносить изменения и уточнения в лексикографическое описание неологической единицы. Таким образом, цифровой неографический ресурс имеет значительные функциональные и структурные преимущества перед полиграфическим вариантом представления словарной информации.

Цифровая неография позволяет своевременно фиксировать языковые изменения, возникающие в современном медиaprостранстве. Динамичное отражение в неологической базе граммати-



ческих, лексико-семантических, стилистических характеристик, парадигматических и синтагматических связей новых языковых фактов позволяет своевременно оценить изменения степени их значимости в коллективном языковом сознании; осознать, в каких значениях они функционируют в речи, спрогнозировать потенциальную закрепленность в лексико-семантической системе.

В цифровом лексикографическом ресурсе «Новое в русской лексике. Словарные материалы» (neolex.iling.spb.ru) использованы основные методики создания ежегодников – «книжек-тетрадей, содержащих отсутствующие в лексикографических изданиях новые слова, новые значения слов и выражения, зарегистрированные по текстам массовых периодических изданий в определенный период времени» [10]. Особенностью словарей-ежегодников является более широкий допуск включаемой лексики по сравнению с нормативным толковым словарем: расширение материала подчинено задачам объективного отражения лексической динамики, выявлению тенденций развития словообразовательной системы русского языка, фиксации изменений русской языковой картины мира.

Процесс лексикографической обработки информации о неологической единице является сложным, многоступенчатым, часто дискуссионным. Особую трудность представляет лексикографическое представление новых заимствованных лексических единиц, которые А.В. Щетина назвала словами с пограничным статусом [11], охарактеризовав таким образом их промежуточное положение между именем собственным и апеллятивом.

К таким словам относятся недавно зафиксированные в русском языке слова чат-джитипи, джитипи, GTP-чат, чатжтп, чатГПТ и некот. другие, находящиеся на стадии узуализации. Задачей статьи является выявление качественного и количественного многообразия вариантов новых заимствованных единиц, связанных с названием нейросети ChatGTP, анализ типов вариативности, отражаемых в медиадискурсе, выбор конкретного объекта (объектов) лексикографического описания, а также разработка и уточнение параметров электронной словарной статьи.

Сложность лексикографирования обусловлена несколькими факторами: терминологическая природа аббревиатуры GTP, словообразовательная структура заимствуемого имени собственного (сращение), различие артикуляционных баз русского и английского языков, наличие большого количества графических и орфографических вариантов, в том числе противоречащих тенденциям орфографической кодификации, неизменяемость компонента джитипи, пограничный характер заимствования (между именем собственным и апеллятивом).

Методология и методика исследования

Методология исследования заключается в последовательном применении разных методов в зависимости от этапа. Во время проведения плановой неологической выборки была сформулирована гипотеза пограничного статусе новых заимствованных слов чатGTP, чат-джитипи и др., в ходе дополнительного мониторинга из базы данных «Интегрум» были извлечены другие варианты анализируемой номинации. С целью оценки частоты употребления выбранной лексики использовались статистические инструменты базы данных. Помимо традиционных корпусных технологий для дополнительного целенаправленного поиска вариантов был использован чат-бот «Поиск в Телеграм», а также инструменты сервиса для оценки пользовательского интереса к конкретным тематикам «Подбор слов» (Яндекс). В дополнении к простому поиску использовались методики включенного наблюдения, а также контент-анализа тематических публикаций медиаресурса «VC.RU» (<https://vc.ru/>).

На заключительном этапе исследования использовался метод опроса в социальной сети ВКонтакте и в мессенджере Телеграм.

Лексическое значение слова определяется набором признаков, обусловленных его парадигматическим противопоставлением какому-то другому слову или нескольким словам и определяющих синтагматическую совместимость слова с его потенциальным контекстным окружением

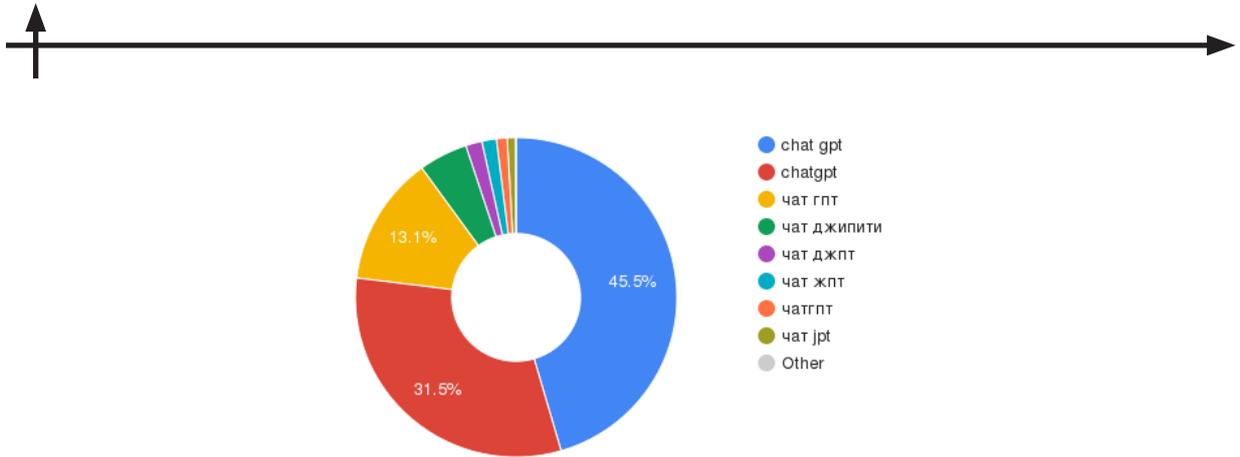


Рис. 2. Сравнение частоты употреблений вариантов названий нейросети в сервисе Яндекса «Подбор слов»
 Fig. 2. Comparison of the frequency of use of neural network name variants in the Yandex “Word Selection” service

[12]. Лингвистическое описание значений, грамматических и стилистических свойств, прагматических характеристик неологических единиц в электронном ресурсе neolex.iling.spb.ru основывается на методиках контекстуального и компонентного анализа, традиционно используемых в лексикографической практике.

Результаты исследования

Появление чат-бота GPT с широкими возможностями (написание текстов, перевод, ведение диалога, создание изображений) привело к активному обсуждению нейросети и связанных с ней моделей в современных медиа. Стремление «русифицировать» сложное название привело к появлению множества вариантов обозначения самой сети и метонимического обозначения используемой ею модели. Разнообразие номинаций было сразу отмечено лингвистами: «...распространенность одной из сетей (ChatGPT) обуславливает появление таких единиц, как GPT-чат, ЧАТЖПТ, Чатжепете и игровых трансформ «Чат Гопоты, он же ГОПОЧАТ». При этом далеко не каждому употребляющему название сети известно, что GPT расшифровывается как Generative Pretrained Transformer (трансформер, обученный для генерации текста)» [2].

Сервис Яндекса «Подбор слов» отражает число показов слова в месяц с учетом параметра слитно/раздельно, без учета регистра. Почти половина запросов в сети выполняются в два слова на английском языке: chat gpt (334239 раз, 45,5%), далее следует запрос-сращение, графически повторяющий название-сращение chatgpt (231127; 31,5%). Среди «кириллических» запросов наиболее частотны словосочетания чат гпт (96092; 13,1%) и чат джитипи (36077; 4,9%). Процент остальных названий колеблется от 0,8 (чат jpt) до 1,6 (чат джпт). Слитное написание чатгпт (662 запроса) программа построения графиков отнесла в сектор «others».

Анализ поисковых запросов позволяет получить приблизительную картину вариантов использования слова и отражает сложный характер первого этапа адаптации заимствования. Эта сложность обусловлена формой заимствуемого слова (имя собственное, оформленное в языке-источнике как сращение), его фонетической структурой с 4 согласными подряд. Кроме того, один из элементов сращения – chat/чат – достаточно освоен русским языком (первая письменная фиксация относится к 1999 г.), осознается и легко выделяется носителями языка, о чем говорит наличие отдельных вариантов написания.

Процесс освоения названия распространенной и набирающей популярность нейросети – яркое явление в современной русской речевой практике. Полученные данные показывают, что распространенность нейросети ChatGPT и связанных с ней технологий, моделей и алгоритмов обусловила востребованность номинации в современной русской речи. В современной русской письменной речи нередки случаи использования неологической единицы в функции



апеллятива. Одним из признаков перехода имени собственного в разряд нарицательных является вариантность (наряду с синонимией и неоднозначностью) [13–15].

Для обсуждения перспектив лексикографирования новейшего заимствования существенным является принятое в современной теории журналистики понятие «качественная пресса»: тип периодических изданий, в которых на высоком профессиональном уровне (взвешенно, аргументированно, нейтрально) освещаются актуальные вопросы политической, экономической, социальной и культурной жизни общества. К таким изданиям специалисты относят «качественные печатные издания федерального охвата с онлайн-версиями (газеты «Аргументы и факты», «Ведомости», «Известия», «Коммерсант», «Комсомольская правда», «Московский комсомолец», «Независимая газета», «Российская газета»; журналы «Огонек», «Профиль», «Эксперт»)» [16].

В процессе анализа реальных словоупотреблений и создания рабочих конкордансов обнаружено, что в качественной прессе на момент написания статьи используется только оригинальное имя собственное в исходной графической форме; не используются (за исключением зоны комментариев) неологические единицы переходного типа: авторы публикаций используют имя собственное ChatGPT, нередко прибегая к различным дискурсивным практикам объяснения (краткое разъяснение, перевод и / или подробный комментарий в скобках, сносках или в отдельном абзаце текста, в т.ч. с графическим выделением). Приведем несколько примеров такого использования.

Антониу Гутерриш заявил, что никто не ожидал, что «искусственный интеллект окажет столь сильное влияние на устойчивое развитие, сферу труда и социальный порядок».

Генсек признался, что «был поражен и впечатлен» радикальным прорывом в возможностях ИИ, который проявился в ChatGPT, а также других подобных сервисах и программах.

ChatGPT – это чат-бот с искусственным интеллектом, разработанный компанией OpenAI. Возможности его максимально широки: чат-бот может вести с пользователями диалоги, отвечать на вопросы, писать код и искать ошибки в нем, разрабатывать бизнес-планы, сочинять стихи и так далее. ChatGPT был запущен 30 ноября 2022 года. Всего за два месяца аудитория активных пользователей достигла 100 млн человек. Сейчас подобные чат-боты разрабатываются и развиваются целым рядом компаний, например Google (Bard), Apple и Baidu (ERNIE Bot).

Выступая на заседании, Антониу Гутерриш обратил внимание на невиданную скорость распространения технологии. GPT-чат достиг 100 млн пользователей всего за два месяца. kommersant.ru 23.07.2023.

В Нью-Йорке появилась новая достопримечательность. Это аппарат, напоминающий одновременно большую стиральную машину, компьютер и еще что-то, вызывающее ассоциации со «Звездными войнами». На передней панели блестят несколько индикаторов. Там же расположена фотокамера, небольшой экран и диск, напоминающий сейфовый замок. С помощью этого диска можно выбрать один из 100 вопросов, появляющихся на экране. От «Мне уехать из Нью-Йорка?» и «Как мне работать лучше?» до «Как мне относиться к "тревожные звонкам в голове?"» и «Не создать ли мне собственную религию?».

После этого пользователю следует ввести дату, время и место своего рождения и ждать. Ответы генерируются технологиями GPT-3 и ChatGPT. kommersant.ru 16.07.2023.

Учёные из Стэнфордского и Калифорнийского университетов провели исследование и пришли к выводу, что виртуальный помощник человека, вокруг которого совсем недавно было столько шума, – чат-бот ChatGPT – быстро деградирует. aif.ru 24.07.2023.

Адаптация заимствованного названия, связанного с важным информационным поводом, происходит «за пределами качественной прессы»: иллюстративный материал, приведенный ниже, извлечен из интернет-источников, не отвечающих требованиям, предъявляемым к охарактеризованному выше типу периодических изданий. Это медиаресурсы, тематически связанные со сферой информационных технологий, новостные источники, форумы и социальные сети.

Анализ функционирования неологических единиц отразил использование фонетико-графических и орфографических вариантов новых заимствований.

Фонетико-графическая вариантность. Многочисленные варианты отражают процессы первого этапа ассимиляции элементов английского сращения. При стремлении к максимально точной передаче происходят замена английской аффрикаты [ʤ] сочетанием ‘джи’ и передача



названий английский букв Р ([pɪ:]) и Т ([ti:]) с помощью русских сочетаний [п'и] и [т'и]. Также возможна передача J, P, T русскими буквами 'г' или 'ж', 'п', 'т': <гэпэтэ> или <жэпэтэ>.

Орфографическая вариантность. В современной русскоязычной письменной речи нередки образования, совмещающие в себе латиницу и кириллицу: *чат jtp*, *GPT-чат*, *чатGPT*, *чат GPT*. Кроме того, встречаются слитные, отдельные и дефисные написания, а также разные варианты использования прописной буквы (например, *чат ДжиПиТи*, *чат-джипити*, *чатгпт*, *чат-ГПТ*, *чатжпт*, *чатЖПТ*, *Чатджипити*, *ЧатДжиПиТи*, *чатджипити*). Можно предположить, что слитное написание, противоречащее орфографической норме, обусловлено двумя факторами: графический облик исходного названия-сращения и характерное для интернет-языка стремление к упрощению. Приведем примеры, отражающие разнообразие вариантов написания названия нейросети в неформальном общении пользователей интернета:

Эти аргументы вызвали активный спор: кто то поддержал такую точку зрения, кто то назвал нейросети инструментом, с помощью которого все же можно создавать настоящее искусство, а кто то раскусил задумку блогера – текст треда был создан при помощи ChatGPT. Я попросил GPT-чат написать мне текст против искусственного интеллекта и в пользу художников.. uarcenter.ru 18.12.2022.

ChatGPT ("Чат ДжиПиТи") – это тоже нейросеть, которая стала доступна для общего пользования в конце 2022 года. vc.ru 05.02.2023.

Бывший студент Российского гуманитарного университета по имени Саша защитил диплом, написанный нейросетью ChatGPT. Он взял чужой диплом про мясокомбинат, перевел его на английский, попросил ЧатЖПТ переделать во что-то похожее на тему «Как управлять игровой компанией», а потом конвертнул обратно из английского в русский. vc.ru 06.02.2023.

ЧатДжиПиТи – чат-бот, способный работать в диалоговом режиме, отвечать на запросы на естественных языках, в том числе на русском. ЧатДжиПиТи обучаем и может понимать «тонкости вопроса», ориентируясь на индивидуальные особенности человека, что делает данный искусственный интеллект «более человечным». ruskline.ru. 14.02.2023.

У ЧатЖеПеТе в России нет шансов. exl.ichost.ru 03.05.2023.

Как бы ни был хорош Чатджипити, у него всё же есть свои ограничения работе, которые негативно сказываются на результат. inttershop.com. 07.06.2023.

Запустил апдейт. Улучшены техника безопасности, саппорт, поддержка. доработан интерфейс. Запихнул в них рецептуры интересных блюд, будут подсказывать во время готовки... Подключил их к Барду и Чатджипити... Чтобы охватывали базу знаний. yablor.ru 19.07.2023.

Заметим, что слитные написания не являются единичными: по статистике медиаресурса «Интегрум», написание чатджипити зафиксировано более 100 раз, а чатгпт – более 500. Несмотря на то, что электронная неологическая база не является нормативной, противоречащие орфографической норме написания в ресурсе не отражаются. Однако в данном конкретном случае – с учетом фактора повторяемости, начальной стадии адаптации неологизма и необычного орфографического представления калькируемой единицы – возможно, целесообразно включение слитных вариантов в словник ресурса (или в справочные зоны основных статей).

Важным критерием перехода имени собственного в нарицательные является написание со строчной буквы: наличие таких вариантов в базах данных является для составителя знаком перехода имени собственного в разряд апеллятивов. На фоне многочисленных вариантов передачи названия нейросети как имени собственного заметны написания вариантов нового слова со строчной буквы: это свидетельствует о том, что в коммуникации специалистов по информационным технологиям осуществляется трансформация английского имени собственного в русское нарицательное существительное.

Сколько же сейчас будет всего на волне чатаджипити) а какой прогресс нам предстоит. vc.ru 07.02.2023.

Искать было лень, поэтому постепенно все запомнилось само собой)) Можно попробовать с чатджипити так). vc.ru 07.03.2023.

Выбираешь, проходишь. ставишь предмет по карте, и проходишь. Пробовал у чатжпт, но это существо выдало такие позорные ответы. biz.mail.ru 24.04.2023.



Изначально письма компании оставались без ответа, что было связано с, возможно, неправильным подходом к деловой коммуникации в контексте культуры и политической ситуации в Турции (абзац отредактирован чатом джипити). vc.ru 18.06.2023.

Все-таки чатгпт и его аналоги не умеют в код. telq.org. 20.07.2023.

А там же новость появилась, дескать, чатгпт отупел. Один и тот же гпт4 в январе давал 97% правильных ответов, а в июне 4%. telq.org 20.07.2023.

Перевод в ChatGPT имеет особенности. Но при этом, чатгпт — инструмент "всё в одном". Мы можем сгенерировать текст по запросу, поправить, перевести, проанализировать и получить код для выкладки в блог с форматированием. habr.com 1.08.2023.

На данный момент чатГПТ — удобный инструмент, но возложить на него судьбы пациентов неправильно и с точки зрения закона, так и этики. habr.com 4.08.2023.

Эти речевые факты свидетельствуют о лексико-семантическом обогащении: в нарицательном значении — как апеллятив — русская калька чат-джипити (чат джипити, чатгпт) используется для обозначения разновидности чат-бота, модели, алгоритма или компьютерной программы с функцией искусственного интеллекта.

Морфологическая (грамматическая) вариантность. Помимо разных вариантов написания не-олексемы, в медиадискурсе наблюдаются разные варианты склонения (*на волне чатаджипити; плагин для создания статей с помощью ЧатаГПТ; все эти ИИ и чатджипити; без ЧатаДжиПити не обойдемся; эти ваши чаты-джипити*), которые отражают процесс грамматической ассимиляции слова. Варианты со склоняемой первой частью свидетельствуют о том, что неологическая единица воспринимается многими как состоящая из двух частей.

Наиболее очевидны признаки апеллятива при употреблении слова в форме множественного числа:

Все эти ИИ и чатджипити — в реальности лишь просто сложные и продвинутое макросы. yablor.ru 29.03.2023.

В итоге получился стартап Potato Parcel.. Вот поистине гениальный продукт, а не эти ваши чаты-джипити. megamozg.ru 28.03.2023 (пример Ю.С. Ридецкой).

Связанность морфологической вариантности, семантического обогащения и варьирования. Пограничный статус неологической единицы определяется содержанием денотативного значения, фактором повторяемости и характером графического облика. Фактор повторяемости обуславливает образование множественного числа существительного, называющего целый класс явлений, и утрату им значения названия конкретной нейросети. В конкорданс вошло некоторое количество контекстов, в которых новое слово обозначает компьютерную программу, модель, алгоритм, инструмент письма, что свидетельствует об изменении объема значения: *«очередное решение, основанное на чатгпт»; «на данный момент чатГПТ удобный инструмент»; «самый обычный чатгпт бот в телеграме».*

Словообразовательная активность и апеллятивность. В конкордансе для цифрового лексико-графического ресурса «Новое в русской лексике. Словарные материалы» (neolex.iling.spb.ru) зафиксированы следующие производные от существительного *джипити*: джипитить, джити-тишный, джипитичный, джипитишка:

Джипитить или не джипитить, вот в чем вопрос. t.me 29.04.2023.

Побежал я к ноутбуку. Нашел помогайку джипитишного этого на русском. Да там и искать особо не пришлось, сразу подвернулся. vk.com 20.03.2023.

Джипитичная девушка. С использованием ChatGPT пользователь разработал виртуальную девушку, способную мыслить и общаться. vk.com 10.05.2023.

Это теперь джипитишка знающая и проанализировавшая ВЕСЬ файл. vc.ru 05.07.2023.

Выявленные речевые факты отражают процессы адаптации неологической единицы в русском языке новейшего периода: на фоне фонетико-графического и орфографического варьи-



рования происходит переход имени собственного в апеллятив: *чат-джипити*, *чат джипити* (существительное с несклоняемым приложением в постпозиции), *чатДжиПиТи*, *чатгпт*, *чатжпт* и некот. др.

Наличие вариантов написания затрудняет процесс лексикографирования, поскольку ни один из вариантов на момент создания электронных словарных статей не кодифицирован (обычно при наличии вариантов написания при выборе заголовочного слова учитываются данные Орфографического академического ресурса «АКАДЕМОС¹», нормативных словарей и правил современной русской орфографии).

С учетом кодификации написания чат-бот, предпочтительным для зоны вокабулы может показаться дефисное написание *чат-джипити*, крайне ограниченно представленное в источниках. Однако в реальной речевой практике преобладают раздельное (*чат джипити*) и слитное написание (*чатджипити*, *чатгпт*, *чатжпт*, а также варианты с прописными буквами во второй части: *чатДжиПиТи*, *чатГПТ*).

Поскольку в источниках зафиксировано несколько вариантов, словарные статьи приводятся к каждому из вариантов, которые связываются между собой при помощи лексикографического инструмента «add орфографический/ орфоэпический вариант».

Неологические единицы, совмещающие латиницу и кириллицу (чатGPT, GPT-чат, GPTчат) включаются в базу без семантической разработки: электронная словарная статья включает зону вокабулы с грамматической справкой и иллюстративный материал с указанием источника.

Толкование к словам терминологического происхождения носит описательно-энциклопедический характер и составляется на базе многочисленных дефиниций, в т.ч. специальных. Несмотря на то, что в медиадискурсе значение слова *чат-джипити* часто объясняется при помощи родового идентификатора *чат-бот*, предпочтительнее представляется более широкое родовое понятие «нейросеть» (совокупность связанных нейронных блоков, выполняющих обработку информации), так как заимствованное слово происходит от названия нейросети.

Кроме того, идентификатор чат-бот ограничивает, «сужает» дефиницию, так как обозначает конкретную программу, выполняющую различные действия по запросу пользователя.

Трудность создания дефиниции обусловлена терминологической природой неологической единицы, специальным характером отражаемого в ней знания. Толкование должно отражать функциональный компонент значения: использование технологии искусственного интеллекта, способность вести коммуникацию, способность создавать тексты разных типов и разных жанров, исправлять ошибки в компьютерных программах.

Словообразовательная справка призвана указать на английское происхождение неологической единицы с переводом входящих в название компонентов (от англ. **chatGPT (Generative Pre-trained Transformer** ‘трансформер, предварительно обученный для генерации текстов’) — *имя собств.*, название нейросети, предназначенной для различных задач обработки естественного языка: генерация текста, машинный перевод, классификация текстов и др.

Поскольку цифровой неологический ресурс является принципиально изменяемым, «активным», в него можно вносить уточнения и по завершении работы над лексикой описываемого года. Динамический характер базы neolex.iling.spb.ru, в которой реализована возможность оперативно изменять данные словарной статьи, позволит в будущем (например, после кодификации одного из вариантов написания) внести коррективы в зоны вокабулы, орфографической и энциклопедической справок.

Заключение

Название нейросети ChatGPT, генетически связанное с терминологией IT-сферы, активно используется в современной русскоязычной медийной сфере.

¹ <http://orfo.ruslang.ru>



В качественной прессе используется название нейросети и алгоритма GPT используются в оригинальном варианте: английское имя собственное. Функционирование слова сопровождается переводом, комментариями, разъяснением; такие дискурсивные практики говорят о первом этапе освоения неологической единицы.

В медиапространстве за пределами качественной прессы наблюдаются признаки перехода имени собственное в разряд апеллятивов, признаками пограничного статуса неологической единицы являются написание со строчной буквы и вариативность.

С учетом значимости явления и обозначающих его слов в жизни современного общества необходимо максимально полно и точно отразить в цифровом ресурсе неологические единицы, связанные с программой ChatGPT и вошедшие в русский язык в 2022–23 гг. Имя собственное (название сети) не подлежит лексикографическому описанию, в словник включаются только апеллятивы.

Неологический цифровой ресурс не является нормативным; в процессе лексикографического описания новых единиц составители ориентируются на действующие орфографические нормы. Поскольку в конкорданс вошло множество вариантов (выбор из которых не представляется орфографически однозначным), представляется необходимым создание электронных словарных статей к каждому из вариантов: раздельное, слитное написание и написание через дефис с учетом различий в передаче английского названия в русской графике. Отказ от лексикографирования таких написаний вступает в противоречие с реальной языковой практикой, поскольку количество употреблений каждого варианта достаточно велико.

Описание графических и орфографических вариантов, отражающие попытки коллективного языкового сознания адаптировать заимствованное слово, представляется очень важным в аспекте изучения языковых контактов, русской языковой картины мира, а также в целях научного анализа функционирования новых заимствованных слов терминологической природы слова в различных типах дискурсов.

СПИСОК ИСТОЧНИКОВ

1. **Козловская Н.В., Янурик С.** ИИ-компози́ты как объект неологии и неографии XXI века / Н. В. Козловская, С. Янурик // *Филологические науки. Научные доклады высшей школы*. 2021. № 2. С. 23–30. DOI: 10.20339/PhS.2-21.023
2. **Плотникова А.М.** Нейросеть как ключевое слово текущего момента // *Филологический класс*. 2023. Т. 28, № 2. С. 45–54.
3. *Лексикография цифровой эпохи: сборник материалов Международного симпозиума (24–25 сентября 2021 г.)* / отв. ред. Е.А. Юрина, С.С. Земичева. Томск: Издательство Томского государственного университета, 2021. 420 с.
4. *Новые слова и словари новых слов 2020: сб. науч. ст.* / отв. ред. Н. В. Козловская. СПб.: ИЛИ РАН. 2020, 220 с.
5. **Priemysheva M.N., Ridetskaya Y.S., Kovalenko K.I.** The Russian academic neography information retrieval resource, *Proceedings of Electronic Lexicography in the 21st Century Conference*: 6, Sintra, 01–03 октября 2019 года. Sintra, 2019. Pp. 213–222.
6. **Ридецкая Ю.С.** Неография и интернет-лексикография // *Фортуна́товские чтения в Карелии: Сборник докладов международной научной конференции*. В 2-х частях, Петрозаводск, 10–12 сентября 2018 года / Под научной редакцией Н.В. Патроевой. Часть 2. Петрозаводск: Петрозаводский государственный университет, 2018. С. 102–104.
7. **Ридецкая Ю.С.** Неография в Институте лингвистических исследований РАН: история и современность // *История, теория и практика академической лексикографии: Юбилейный сборник научных статей*. Санкт-Петербург: Институт лингвистических исследований РАН, 2022. С. 151–162.



8. **Рябцева Н.К.** Современная лексикография, дополненная лингвистическая реальность и проблемы терминологии // *Лингвистика и методика преподавания иностранных языков*. 2022. № 1 (16). С. 75–110. DOI: 10.37892/2218-1393-2022-16-1-75-110
9. **Чердаков Д.Н.** Словари русского языка в интернете // *Journal of Applied Linguistics and Lexicography*. 2020. Т. 2, № 2. С. 193–209. DOI: 10.33910/2687-0215-2020-2-2-193-209
10. **Котелова Н.З.** Избранные работы / Российская академия наук; Институт лингвистических исследований. СПб.: Нестор-История, 2015. 276 с.
11. **Щетинина А.В.** Между именем собственным и апеллятивом: лексикографирование слов с пограничным статусом (на примере неологизма Брексит) // *Вестник Волгоградского государственного университета. Серия 2: Языкознание*. 2021. Т. 20, № 5. С. 32–42. DOI: 10.15688/jvolsu2.2021.5.3
12. **Уфимцева А.А.** Лексическое значение: Принцип семиологического описания лексики. М.: Наука, 1986. 240 с.
13. **Куслик А.М.** К вопросу о переходе имён собственных в нарицательные // *Учёные записки, ЛГПИ им. А.И. Герцена*. 1959. Т. 189 Вып. 2. С. 155–177.
14. **Леваева Л.В.** Процесс перехода имен собственных в имена нарицательные // *Вестник Волжского университета им. В.Н. Татищева*. 2010. № 4. С. 133–136.
15. **Липустина О.М., Ткаченко Е.Е.** Системные признаки в процессе перехода имен собственных в имена нарицательные // *Евразийский союз ученых*. 2014. № 5-4(5). С. 39–42.
16. **Алевизаки О.Р., Касперова Л.Т., Славкин В.В.** Неологизмы в качественной прессе 2020 года // *Новые слова и словари новых слов. 2020: Сборник статей / отв. ред. Н.В. Козловская, Ин-т лингв. исслед. РАН*. СПб.: ИЛИ РАН, 2020. С. 22–27.

REFERENCES

- [1] **Kozlovskaya N.V., Yanurik S.** “II-Composites” as an object of neology and neography of the XXI century, *Philological sciences. Scientific reports of the Higher School*, 2 (2021) 23–30. DOI: 10.20339/PhS.2-21.023
- [2] **Plotnikova A.M.** Neyroset’ as a Keyword of the Current Moment, *Philological Class*. 2 (28) (2023) 45–54.
- [3] **Yurina E.A., Zemicheva S.S., ed.** *Lexicography of the digital age: proceedings of the International Symposium (September, 24–25, 2021) ed. by Tomsk: TSU Press, 2021.*
- [4] **Kozlovskaya N.V., ed.** *New Words and Dictionaries of New Words. 2020. Collection of Scientific Articles. Saint Petersburg, Institute for Linguistic Studies of the RAS, 2020.*
- [5] **Priemysheva M.N., Ridetskaya Yu.S., Kovalenko K.I.** The Russian academic neography information retrieval resource, *Proceedings of Electronic Lexicography in the 21st Century Conference: 6, Sintra, 01–03 october 2019. Sintra, 2019. Pp. 213–222.*
- [6] **Ridetskaya Yu.S.** Neography and Internet lexicography, *Fortunatov Readings in Karelia: Collection of reports of the international scientific conference. Petrozavodsk, September 10–12, 2018, Ed. by N.V. Patroeva (2). Petrozavodsk: Petrozavodsk State University, 2018. Pp. 102–104.*
- [7] **Ridetskaya Yu.S.** Neography at the Institute for Linguistic Studies of the Russian Academy of Sciences: history and modernity, *History, theory and practice of academic lexicography: Anniversary collection of scientific articles. St. Petersburg: Institute for Linguistic Research RAS, 2022. Pp. 151–162.*
- [8] **Ryabtseva N.K.** Modern lexicography, augmented linguistic reality and problems of terminography, *Linguistics and methods of teaching foreign languages*, 1 (16) (2022) 75–110. DOI: 10.37892/2218-1393-2022-16-1-75-110
- [9] **Cherdakov D.N.** Dictionaries of the Russian language on the Internet, *Journal of Applied Linguistics and Lexicography*, 2 (2) (2020) 193–209. DOI: 10.33910/2687-0215-2020-2-2-193-209
- [10] **Kotelova N.Z.** *Selected Works, Russian Academy of Sciences; Institute for Linguistic Research. St. Petersburg: Nestor-History, 2015.*
- [11] **Shchetinina A.V.** Between proper name and appellative: lexicography of words with borderline status (exemplified by neologism Brexit), *Bulletin of the Volgograd State University. Series 2: Linguistics*, 20 (5) (2021) 32–42. DOI: 10.15688/jvolsu2.2021.5.3



[12] **Ufimtseva A.A.** Lexical Meaning: The Principle of Semiological Lexical Description, M.: Nauka, 1986.

[13] **Kuslik A.M.** On the issue of the transition of proper names into common nouns, Uchenye zapiski, LGPI im. A.I. Herzen, 189 (2) (1959) 155–177.

[14] **Levaeva L.V.** The process of transition of proper names into common nouns, Bulletin of the Volga University after name V.N. Tatishchev, 4 (2010) 133–136.

[15] **Lipustina O.M., Tkachenko E.E.** Systemic features in the process of transition of proper names into common nouns, Eurasian Union of Scientists, 5-4 (5) (2014) 39–42.

[16] **Alevizaki O.R., Kasperova L.T., Slavkin V.V.** Neologisms in high-quality print media 2020, New Words and Dictionaries of New Words. Collection of Scientific Articles. 2020. Saint Petersburg, Institute for Linguistic Studies of the RAS, 2020. Pp. 22–27.

СВЕДЕНИЯ ОБ АВТОРЕ / INFORMATION ABOUT AUTHOR

Козловская Наталия Витальевна

Nataliia V. Kozlovskaja

E-mail: mnegolosbyl@gmail.com

<https://orcid.org/0000-0001-7272-7979>

Поступила: 30.08.2023; Одобрена: 26.10.2023; Принята: 10.11.2023.

Submitted: 30.08.2023; Approved: 26.10.2023; Accepted: 10.11.2023.