

Терминология цифрового мира в свете коммуникативистики

Terminology of the Digital World in the Light of Communication Studies

DOI: 10.12737/2587-9103-2024-13-2-39-43

Получено: 12 января 2024 г. / Одобрено: 30 января 2024 г. / Опубликовано: 26 апреля 2024 г.

**А.П. Миньяр-Белоручева**

Д-р филол. наук, профессор,
Московский государственный
университет имени М.В. Ломоносова,
Россия, 119991, Москва, Ленинские горы, 1,
e-mail: ostvera@mail.ru

A.P. Minyar-Beloroucheva

Professor, Dr. Habil Professor,
Lomonosov Moscow State University,
1, Leninskiye Gori, Moscow, 119991, Russia,
e-mail: ostvera@mail.ru

**П.И. Сергиенко**

Канд. филол. наук, преподаватель,
Московский государственный
университет имени М.В. Ломоносова,
Россия, 119991, Москва, Ленинские горы, 1,
e-mail: poserg@bk.ru

P.I. Sergienko

Lecturer, Ph.D.,
Lomonosov Moscow State University,
1, Leninskiye Gori, Moscow, 119991, Russia,
e-mail: poserg@bk.ru

Аннотация

Развитие цифровых технологий в современном мире, в частности искусственного интеллекта (ИИ) привело к созданию уникальной языковой модели *ChatGPT*, способной общаться с человеком. Данная статья посвящена рассмотрению особенностей функционирования языковой модели *ChatGPT*, являющейся продуктом искусственного интеллекта. Обученная человеком, данная модель способна общаться с пользователями, создавая впечатление полноценной коммуникации, наглядно демонстрируя, что обмен информацией может происходить не только между людьми, но и между человеком и машиной. Коммуникативные сдвиги подобного рода оказались возможны в результате достижений цифровой революции, формирующей новое киберпространство существования человека. Это свидетельствует о радикальных изменениях, которые складываются в обществе, и приводит к смене парадигм и формированию нового мышления и его воплощения в языке. Цель данной статьи заключается в анализе терминов и их дефиниций, раскрывающих сущность и функции языковой модели *ChatGPT*. Задачи определяются поставленной целью и сводятся к ряду операций, необходимых для ее достижения. Новизна данной работы обусловлена материалом исследования, который впервые становится предметом лингвистического анализа. Методы, используемые в предлагаемой статье, такие как семантический анализ терминов *ChatGPT*, их синтез, а также классификация позволили представить терминосистему указанной области в развитии, формирующую новые знания и способствующую эффективной коммуникации. В результате анализа терминов авторам удалось создать четкую картину операции, связанной с обучением *ChatGPT*, выявить основные функции данной языковой модели, проследить пути ее дальнейшего совершенствования.

Ключевые слова: терминология, *ChatGPT*, семантический анализ, функции, новое знание.

Введение

Общепризнанным является восприятие языка как важнейшего средства человеческого общения, который выступает как орудие познания, как инструмент мышления. «Благодаря этому коммуникация между людьми является основным механизмом становления человека как социальной личности <...>» [2, с. 7]. О.Я. Гойхман подчеркивает, что «главная цель речевой коммуникации — обмен информацией различного рода [2, с. 7].

Abstract

The development of digital technologies in the modern world, in particular artificial intelligence (AI), has brought about the creation of *ChatGPT*, a unique *ChatGPT* language model capable of communicating with humans. This article is devoted to the analysis of the functioning features of the *ChatGPT* language model, which is a product of artificial intelligence. Trained by a human, this model is able of communicating with users, creating the impression of full-fledged communication, vividly demonstrating that information exchange can occur not only between people, but also between a person and a machine. Communicative shifts of this kind have become possible as a result of the achievements of the digital revolution, which is shaping a new cyberspace of human existence. This indicates the radical changes that are taking place in society, and lead to a paradigm shift and the formation of new thinking and its embodiment in language. The purpose of this article is to analyze the terminology and its definitions, revealing the essence and functions of the *ChatGPT* language model. The tasks of this paper are defined by the purpose set for the investigation and boil down to a number of operations necessary to achieve it. The novelty of this work is due to the research material, which has become the subject of linguistic analysis for the first time. The methods used in the article, such as the semantic analysis of *ChatGPT* terminology, its synthesis, as well as classification, has made it possible to present the terminological system of this field in the process of creating new knowledge and contributing to effective communication. As a result of the analysis of the language units in question, the authors have managed to understand the operation associated with *ChatGPT* training, identifying the main functions of this language model, and tracing the ways of its further development.

Keywords: terminology, *ChatGPT*, semantic analysis, functions, new knowledge.

В настоящее время общение и обмен информацией осуществляется не только между людьми с использованием естественного языка. Развитие современных цифровых технологий позволяет осуществлять коммуникацию между человеком и машиной, что приводит к коммуникативным сдвигам, которые складываются в обществе, обуславливая смену парадигм.

Современное общение между человеком и машиной вызвано нарождающимся цифровым миром,

создающим особую среду, во многом обеспеченную традиционными моделями человеческого языка, или языковым паттернами. Колоссальные трансформации происходят в реальном и виртуальном пространстве. У человека, ощущающего себя частью компьютерной реальности, происходит изменение концептуальной картины мира и ценностных установок, чему во многом способствуют новые терминологические единицы, отражающие преобразования, происходящие в материальном и ментальном пространстве современных людей.

Новый вид коммуникации осуществляется в виртуально-реальном мире, что оказывается маркером современной эпохи. Искусственно созданный цифровой мир становится неотъемлемой частью реального существования человечества, что приводит к смене ценностных ориентиров, особенно четко проявляющихся в новой культуре западного общества. Новые ценности распространяются по всей планете в режиме реального времени во многом благодаря цифровым технологиям [3] и английскому языку, являющимся *lingua franca* современности.

Состояние дел в изучаемой области

Как отмечалось, в настоящее время новый вид коммуникации обеспечивается достижением человека в области искусственного интеллекта посредством нейронных сетей. Созданные по образцу работы человеческого мозга нейронные сети используются для анализа больших данных — *Big Data*, обработки информации и выполнения многочисленных когнитивных задач. Нейронные сети, разработанные человеком, являются мощным инструментом, необходимым для моделирования способов, которыми люди переосмысливают и понимают язык для осуществления коммуникации. На современном этапе изучение естественного языка в его взаимодействии с мышлением человека получило новое значение для раскрытия современных горизонтов научного знания и путей эволюции человеческого общества, важнейшим вектором развития которого является слияние терминологии нейронных сетей с концепциями современной коммуникативистики. Особенность нейросетей состоит в их способности понимать язык и мышление, т.е. общаться с человеком в непривычной для него форме. Человек не видит невербальные реакции на свои высказывания и не может оценить сенсорные механизмы восприятия собственных слов, что приводит к изменению его ценностной картины мира. Рассмотрение основных факторов, влияющих на изменение ценностной картины мира, основывается на выявлении закономерностей развития общества и их воплощении в языке. Лексические единицы компьютерной сферы, тиражируемые в совре-

менном мировом языковом пространстве, отражают новые ценности и пути развития общества. В то же время, цифровая революция, принеся с собой колоссальный информационный объем, вербализованный как *Big Data*, нивелировала существующие территориальные границы, сжав расстояния и увеличив скорость распространения новых идей.

Результат и дискуссия

Цифровые технологии, ставшие доступными человечеству, вводят людей в виртуальную реальность неизведанной территории киберпространства, способной модифицировать современную жизнь. Истоками стремления человека к познанию и покорению нового измерения его ментального бытия, способного усовершенствовать его как физические, так и когнитивные способности, являются древние мифы о сверхчеловеке, которые имеются у каждого народа. В идеальном представлении компьютерные чипы, встраиваемые в организм человека, наделяют его сверхспособностями, благодаря которым он становится умнее, быстрее, сильнее, креативнее, независимо от происхождения и социального статуса. Следует отметить, что во всех достижениях человека особую роль играет язык, задающий новые параметры идеального человека будущего. Данное положение объясняет повышенный интерес лингвистов, направленный на изучение языкового воздействия на мышление человека и перенос, а также применение полученных знаний в новой технологической среде, открывшей возможности для новейшей технологии пространства искусственного интеллекта. Это поднимает вопрос о бытии в реальном мире и виртуальной среде, который получил новое звучание в условиях объединения материального и иллюзорного мира киберпространства, с учетом возможности программирования моделей искусственного интеллекта на самообучение (*Deep Learning*) и генерирование текстов с их последующим использованием в коммуникации человек-машина.

Поскольку в настоящее время происходит переход в виртуальное пространство компьютерной реальности, повышенное внимание уделяется взаимосвязи языка и мышления с позиций исследования когнитивных процессов, происходящих в сознании человека с учетом их дальнейшего перенесения в область компьютерных технологий, что приводит к модернизации мира в режиме реального времени. Когнитология предоставляет возможность не только осуществить категоризацию окружающего мира [1], но и рассмотреть взаимосвязь между терминологией нейронных сетей и концепцией когнитивной лингвистики, позволяющей систематизировать знание о моделях и функциях новой цифровой среды. Тандем,

объединяющий исследования терминологии математических моделей нейронных сетей как системы и компьютерных технологий позволяет выявить механизмы работы языкового мышления и путей вербализации компьютерного знания, что способствует развитию нового вида коммуникации между машиной и человеком. В связи с тем, что приоритет изучения нейронных сетей находится в области цифрового дискурса английского языка, термины, отражающие появление и функционирование новых сфер знания [4], создаются на английском языке и в процессе коммуникации транслируются на другие языки по всему миру. Изучение англоязычной терминологии в области компьютерной технологии обеспечивает наиболее широкий взгляд на рассматриваемую проблематику, поскольку терминологические единицы профессионального языка свидетельствуют о новых достижениях в области знания, способствующего осуществлению нового вида коммуникации.

Основным инструментом общения между человеком и машиной в настоящее время является *Chat Bot GPT*, впервые разработанный американской компанией *Open AI*, расположенной в Сан-Франциско. *ChatGPT (Generative Pre-trained Transformer* — генеративный предварительно обученный трансформер) представляет собой модель генерирования текста, способную, общаясь с человеком, понимать его вопросы, реагировать на них, давать советы и комментарии, выражая при этом, как показывает язык, чувства сенсорного восприятия информации.

Общение человека с *ChatGPT*, основным продуктом искусственного интеллекта (ИИ), свидетельствует о вступлении мира в новую эпоху онтологической революции. Терминологические единицы, характеризующие новую модель ИИ, способную общаться с человеком, демонстрируют возможность машины вовлекать его в виртуальный мир, делая его неотъемлемой частью пространства компьютерных нейронных сетей.

При работе чат-бота *GPT* как модели коммуникации, обладающей своей терминологией, отражающей его функции и возможности, во главу угла поставлены язык человека и его мышление, при этом базовые термины, характеризующие лингвистическую модель коммуникации с человеком, соответственно связаны с этими областями. Исследователи отмечают, что терминосистема профессиональной области «строится на ключевых понятиях, и их описанию нужно уделить наибольшее внимание, тем более что система формируется и как совокупность единиц, и как совокупность очень сложной сетки знаний, которая может быть представлена когнитивными картами, сценарными планами и т.д.» [5, с.139]. Использование взаимосвязи терминологии компью-

терных нейронных сетей и достижений в области коммуникативистики приводит к пониманию того, как математические модели ИИ могут взаимодействовать с человеком. Другими словами, наиболее важным оказывается исследование новых терминов в системе в рамках означенной области, что по сути является способом передачи знания и возможностью обобщения и умножения научных достижений.

Благодаря исследованию терминосистем, которые репрезентируют работу *ChatGPT*, исследователи имеют возможность получить знание о возможностях и функциях, выполняемых генеративной моделью искусственного интеллекта, что необходимо для дальнейшего усовершенствования коммуникации между человеком и машиной. Как правило, в дефинициях терминов заложена информация, раскрывающая функции и характеристики новых предметов, моделей и понятий. В связи с отсутствием официальных лексикографических источников по данной проблематике, для настоящего исследования здесь и далее авторы рассматривают определения, предложенные компьютерными руководствами и тематическими глоссариями, размещенными в глобальной сети Интернет [6].

Представляется необходимым начать анализ сущности языковой модели ИИ с определения термина *ChatGPT* (чат *GPT*). Согласно существующим словарям, под *ChatGPT* понимается программа, выясняющая и удовлетворяющая потребности пользователя компьютерной системы, которая была усовершенствована и доведена до возможности общения с человеком. Лингвистическая функция данного термина раскрывается в его дефиниции, которая представлена как языковая модель, обученная генерировать текст на основе речевых образцов, извлеченных из большого количества примеров: *ChatGPT — a type of machine learning model that can generate text based on the patterns it has learned from a large dataset of examples.*

Отметим, что для классификации терминов, отражающих общение новой языковой модели с человеком, профессиональные языковые единицы необходимо подразделить на несколько групп.

К нулевой группе — *pre-training* — можно отнести термины, касающиеся анализа большого объема текстов с целью извлечения лексического и грамматического материала, необходимого для машинного обучения языковой модели, о чем свидетельствует термин *pre-training*. Данный термин показывает, что для успешного обучения языковой модели ИИ *ChatGPT* общению с человеком, потребовалось изучить закономерности структурирования и функционирования естественного языка на большом объеме текстов: *The process of training a language model on a large dataset to learn the patterns of natural language.*

Особой задачей для осуществления эффективной коммуникации человека и машины является преодоление сложности коллокации и коллигации, то есть выстраивания лексических единиц и синтаксических структур, обозначенное термином *perplexity*. Дефиниция данного термина показывает, насколько хорошо языковая модель может составлять фразы и предложения — *perplexity is a measure of how well a language model can predict a sequence of words*. Следует отметить, что обучение языковой модели английскому языку стало возможным после извлечения языкового материала из большого объема текстов и его классификации.

В первую группу целесообразно включить термины, касающиеся обучения общению *ChatGPT* с человеком, а именно: обучение с нуля — *zero-short learning*, краткосрочное обучение — *few-short learning*, самообучение — *meta-learning*, совершенствование и оптимизация работы — *learning rate*, *backpropagation*, а также переобучение — *overfitting and regulation*. Для создания *ChatGPT* потребовалось не только изучить языковой материал, но и обучить машину самосовершенствоваться и саморазвиваться, о чем свидетельствует анализ терминосистемы первой группы терминов, касающихся подготовки машины к коммуникации с человеком.

К наиболее важным терминам данной группы следует отнести лексические единицы, связанные с различными этапами подготовки к коммуникации человека и машины. Мы опускаем технические особенности создания лингвистического продукта ИИ и сосредотачиваемся на вопросе, с которым столкнулись разработчики данной модели — «чему» необходимо научить машину для успешной коммуникации с человеком. Это потребовало постановки и осуществления ряда задач, которые сводились к проведению большой подготовительной работы. В первую очередь создателям *ChatGPT* необходимо было на структурном уровне детально изучить естественный английский язык и обучить языковую модель извлеченным лексическим и грамматическим структурам, а также объяснить, что из себя представляет языковой продукт нейронных сетей. Все это нашло отражение в соответствующих *ChatGPT*-терминах и их дефинициях.

На начальном этапе происходит обучение машины с нуля без четкого определения поставленной задачи: *Zero-shot learning is the ability of a language model to perform a task without being explicitly trained on that task*. Данный термин демонстрирует отсутствие целенаправленной задачи и выстраивания обучения по ходу работы, поскольку в данной языковой модели *ChatGPT* априори была заложена способность к обучению, т.е. метаобучения — *meta-learning* — the

ability of a language model to learn. Для реализации самообучения были выделены несколько параметров, что проявляется в соответствующих терминах — *BERT*, *RoBERTa* — параметры, указывающие на степень способности языковой модели к самообучению.

О способности *ChatGPT* к переобучению свидетельствует термин *overfitting*, означающий явление, когда, ввиду несовершенства обработки текстового материала, языковая модель усваивает отдельные слова, вместо клишированных фраз и готовых синтаксических конструкций — *a phenomenon in machine learning where a model performs well on the training data, but poorly on the test data, due to memorizing the training data instead of learning generalizable patterns*.

Ко второй группе относятся термины, дефинирующие порождение текста языковой моделью, что является основной задачей в процессе коммуникации *ChatGPT* с человеком. Термином *text generation* обозначено генерирование текста языковой моделью на основе заданной темы или контекста — *the process of generating new text based on a given prompt or context*. Дефиниция термина *fine-tuning* показывает, что для того, чтобы обученная языку модель *ChatGPT* могла общаться с человеком в рамках профессионального и бытового дискурсов, ее необходимо было адаптировать к переходу на уровень речи — *the process of adapting a pre-trained language model to a specific task or domain*. Данный термин и его дефиниция свидетельствуют от том, что при переходе от знания системы языка к его реальному употреблению в речи требуется практика и время.

К третьей группе относятся термины *ChatGPT*, в дефинициях которых отражены функции, выполняемые данной языковой моделью при коммуникации с человеком, которые можно подразделить на несколько групп, в частности, можно выделить функцию анализа языка (*Transformer Architecture*, включающую *encoder* и *decoder*), функцию анализа текста (*NER* — *named entity recognition*), коммуникативную функцию (*conversational AI*) *per se*, включающей функцию обратной связи с человеком (*Human-in-the-Loop (HITL)*).

Особого внимания заслуживает функция выявления эмоциональности высказывания (*sentiment analysis*, а также функция, обеспечивающая процесс сохранения информации в памяти языковой модели (*long short-term memory*). Так, термином *sentiment analysis* обозначается способность чат-бота *GPT* осуществлять анализ эмоциональной составляющей, заложенной в тексте — *the process of determining the emotional tone of a piece of text*. Также к данной группе относятся термины, называющие чувства и элементы сенсорного восприятия текста, механизм внимания — *attention mechanism*, *muly-head attention*. Данный

компонент структуры языковой модели позволяет фокусироваться на разных частях входящего текста посредством его декодирования на знаки, приемлемые для восприятия *ChatGPT* — *a component of the transformer architecture that allows the model to focus on different parts of the input text*. Термины, выявляющие функции языковой модели, свидетельствуют об их многообразии, способствующем успешной коммуникации человека с нейросетью в новых условиях цифровой среды. Введение цифровых технологий во всех сферы деятельности современного общества неминуемо приведет к созданию нового образа жизни человека новой эпохи, неотъемлемой частью которого становится искусственный интеллект, который уже способен понимать мышление человека, выраженного в языке.

Заключение

Подводя итог, следует отметить, что было разработано четыре поколения языковой модели *ChatGPT* от компании *Open AI* и ожидается выход чата пятого поколения. Модели *ChatGPT* постоянно модернизи-

руются, в разы увеличивая количество параметров, на которые опирается ИИ. Становясь более совершенными, они приобретают все больший набор функций, необходимых для полноценного общения с человеком. Анализ терминосистемы языковой модели *ChatGPT* раскрывает многообразие функций и параметров новых продуктов от искусственного интеллекта. Профессиональные термины играют существенную роль в передаче большого количества информации, способствуют появлению нового знания и оказывают существенное влияние на коммуникацию. Уникальная языковая модель *ChatGPT*, способная общаться с человеком, открывает новые возможности для существования человека в формирующейся среде глобального киберпространства. Терминосистема, передающая совокупность знаний о новых достижениях человека в области цифровых технологий, выполняя свое значение и способствуя созданию новой концептуальной картины мира человека, подготавливает его к жизни во взаимосвязанном пространстве виртуального и реального мира.

Литература

1. Болдырев Н.Н. Когнитивная семантика. Введение в когнитивную лингвистику [Текст]: курс лекций / Н.Н. Болдырев. — 5-е изд., испр. и доп. — Тамбов: Державинский, 2021. — 236 с.
2. Гойхман О.Я. Речевая коммуникация [Текст]: учебник / О.Я. Гойхман, Т.М. Надеина. — М.: ИНФРА-М, 2003. — 272 с.
3. Гончарова Л.М. Профессиональная коммуникация в ракурсе вопросов современной коммуникативистики: вызовы и задачи в условиях цифровизации и пандемии [Текст] / Л.М. Гончарова // Научные исследования и разработки. Современная коммуникативистика. — 2021. — Т. 10. — № 3. — С. 5–10.
4. Кубрякова Е.С. Язык и знание: на пути получения знаний о языке: части речи с когнитивной точки зрения. Роль языка в познании мира [Текст] / Е.С. Кубрякова. — М.: Яз. славян. культуры, 2004. — 555 с.
5. Новодранова В.Ф. Проблемы терминологического образования в когнитивно-коммуникативном аспекте [Текст] / В.Ф. Новодранова // Лексикология. Терминоведение. Стилистика: сб. науч. тр., посвящ. юбилею В.М. Лейчика. — М.: Рязань, 2003. — С. 150–155.
6. ChatGPT and GPT Glossary. URL: <https://trickmenot.ai/chatgpt-and-gpt-glossary-terms> (accessed 09.01.2024).

References

1. Boldyrev N.N. Kognitivnaya semantika. Vvedenie v kognitivnuyu lingvistiku: kurs lekcij [Cognitive semantics. Introduction to cognitive linguistics: a course of lectures]. 5-e izd., ispr. i dop. Tambov: Derzhavinskij, 2021. 236 s.
2. Goykhman O.Ya., Nadeina T.M. Rechevaya kommunikaciya: Uchebnik [Speech communication: Textbook]. M.: INFRA-M, 2003. 272 s.
3. Goncharova L. M. Professional`naya kommunikaciya v rakurse voprosov sovremennoj kommunikativistiki: vy`zovy` i zadachi v usloviyax cifrovizacii i pandemii [Professional communication in the perspective of modern communication studies: challenges and tasks in the context of digitalization and a pandemic]. Nauchny`e issledovaniya i razrabotki. Sovremennaya kommunikativistika [Scientific research and development. Modern communication science]. 2021, v. 10, i. 3, pp. 5–10.
4. Kubryakova E.S. Yazy`k i znanie: na puti polucheniya znaniy o yazy`ke: chasti rechi s kognitivnoj tochki zreniya. Rol` yazy`ka v poznanii mira [Language and knowledge: on the way to gaining knowledge about language: parts of speech from a cognitive point of view. The role of language in understanding the world]. M.: Yaz. slavyan. kul`tury`, 2004. 555 s.
5. Novodranova V.F. Problemy` terminoobrazovaniya v kognitivno-kommunikativnom aspekte [Problems of terminology formation in the cognitive and communicative aspect] / V.F. Novodranova // Leksikologiya. Terminovedenie. Stilistika: sb. nauch. tr., posvyashh. yubileyu V.M. Lejchika. M.: Ryazan`, 2003. S. 150–155
6. ChatGPT and GPT Glossary, URL: <https://trickmenot.ai/chatgpt-and-gpt-glossary-terms> (accessed 09.01.2024).