



DOI: 10.22363/2313-2299-2024-15-2-364-375

EDN: PNJPNY

УДК [811.161.1:811.581]’373:004.8

Научная статья / Research article

Цифровой страт парадигмы языка

К.П. Жикулина  , Н.В. Перфильева , Мань Ли

Российский университет дружбы народов, Москва, Российская Федерация

 zhikulina_kp@pfur.ru

Аннотация. Специфика функционирования искусственного интеллекта и цифрового языка в современном мире является актуальным направлением междисциплинарных, в том числе, лингвистических исследований. Предмет анализа — голосовой помощник Алиса от компании Яндекс со встроенной нейросетью Yandex GPT2. Объект исследования — функционирование ИИ на материале фразеологических единиц и устойчивых выражений китайского и русского языков с универсальным семантическим компонентом *смерть*. Концептуальная категория *смерть* выбрана потому, что в русской и китайской культурной традиции существует богатая база, восходящая к фольклорным образам. Новизна исследования заключается в том, что понятие цифрового страта как отдельного компонента языковой парадигмы, по Н.И. Толстому, вводится и рассматривается впервые. Выделено 4 типа генерации на слова-стимулы в ИИ: 1) отсутствие генерации; 2) генерация существующих фразеологизмов и устойчивых выражений с прямой и непрямой номинацией; 3) гибридная генерация; 4) цифровая или произвольная генерация. Методом сплошной выборки было получено 50% генераций ИИ на русском и на переводном материале. Результаты исследования показали, что происходит формирование цифрового страта парадигмы языка за счет развития виртуальной среды языкового функционирования. Со временем текстовые генерации, которые в настоящее время представляются алогичными, могут войти в живой разговорный стиль и стать общеупотребительными за счет массового использования голосовых систем в различных областях жизнедеятельности.

Ключевые слова: искусственный интеллект, искусственная среда языкового функционирования, цифровой язык, голосовой помощник Алиса, машинный фонд русского языка (МФРЯ), фразеологические единицы, устойчивые выражения, китайский язык

Финансирование. Благодарности:

Публикация выполнена в рамках грантовой поддержки научных проектов РУДН 124022500238–1 «Модель мультязычного терминологического словаря».

История статьи:

Дата поступления: 01.02.2024

Дата приема в печать: 15.02.2024

© Жикулина К.П., Перфильева Н.В., Ли Мань, 2024

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

Для цитирования:

Жикулина К.П., Перфильева Н.В., Ли Мань. Цифровой страт парадигмы языка // Вестник Российского университета дружбы народов. Серия: Теория языка. Семиотика. Семантика. 2024. Т. 15. № 2. С. 364–375. <https://doi.org/10.22363/2313-2299-2024-15-2-364-375>

The Digital Stratum of a Linguistic Paradigm

Christina P. Zhikulina  , Natalia V. Perfilieva , Man Li

RUDN University, Moscow, Russian Federation

 zhikulina_kp@pfur.ru

Abstract. The specific of functioning for artificial intelligence and digital stratum is an important direction of interdisciplinary, including linguistics, researches in the modern world. The subject of the study is the voice assistant Alice of the Yandex company with the inbuilt Yandex GPT2 neural network. The object under the study is the AI functioning on the materials of phraseological units and set phrases of the Chinese and Russian languages with the universal semantic component *смерть* / *death*. The conceptual category of *смерть* / *death* was chosen due to the existence of the Russian and Chinese cultural tradition represented on the rich background ascendant to folklore images. The novelty of the study consists in understanding the notion of “a digital stratum as a separate component of the linguistic paradigm, according to N.I. Tolstoy, and it is introduced and discussed for the first time. In course of linguistic experiment, we managed to reveal four types of generations on the AI stimulus words: 1) the absence of generation; 2) the generation of the existing phraseological units and set phrases of direct and indirect nominations; 3) the hybrid generation; 4) the digital or arbitrary generation. By means of the continuous sampling method there were identified 50 % of the AI generations both in Russian and in translated into Russian materials. The results of the study allow conclude that there takes place of forming the digital stratum of the linguistic paradigm owing to the evolution of the virtual environment of language functioning. With time, textual generations which at present seem to be antilogical, could enter the live conversational style and become widely-used ones due to the mass use of voice systems in various spheres of daily living activities.

Keywords: artificial intelligent (AI), artificial field of language operations, digital language, voice assistant Alice, the machine fund of the Russian language (MFRL), phraseological units, set phrases, Chinese

Financing. Acknowledgements:

The publication has been supported by the RUDN University Scientific Projects Grant System, project No 124022500238–1 “Multilingual terminological dictionary model”.

Article history:

Received: 01.02.2024

Accepted: 15.02.2024

For citation:

Zhikulina, Ch.P., Perfilieva, N.V. & Li, Man (2024). The Digital Stratum of a Linguistic Paradigm. *RUDN Journal of Language Studies, Semiotics and Semantics*, 15(2), 364–375. <https://doi.org/10.22363/2313-2299-2024-15-2-364-375>

Введение

Необходимость применения электронной вычислительной машины (далее — ЭВМ) в лингвистике возникла в Советском Союзе в XX в. Производство тех времен предъявляло много требований к использованию автоматизированных систем типа *человек-машина* и взаимодействию гуманитарных наук с естественными [1. С. 3]. В этой связи стала актуальной задача расширения базы многоаспектных исследований русского языка с целью обеспечения большего круга пользователей к новым разработкам.

Потребность применения ЭВМ в сфере автоматической обработки информации на естественном языке привела к созданию «Машинного фонда русского языка» (далее — МФРЯ) [1. С. 4], идея создания которого принадлежала академику А.П. Ершову: «любой прогресс в области построения моделей и алгоритмов останется <...> академическим упражнением, если не будет решена наиважнейшая задача создания Машинного фонда русского языка», решение которой будет иметь «большую научную, общекультурную и прикладную ценность» [2. С. 14]. В то время было опубликовано работы, сохраняющие свою актуальность и сегодня, обсуждающие вопросы архитектуры и концепции МФРЯ [3]; автоматизации лексикографических работ и результатов применения ЭВМ [4]; проектов создания Диалектологического текстового подфонда МФРЯ [5]; состояния и перспектив автоматического словаря народного говора [6] и др.

Машинный фонд — это лингвистический резерв, который в настоящее время поддерживает технологии искусственного интеллекта (далее — ИИ) и затрагивает сферы технологического мира от электронных словарей до банковских услуг. *Машине* необходимо обрабатывать огромные пласты информации, которую ИИ получает из поисковых систем Интернета, однако есть примеры, когда для генерации ответов *машиной* также используется Национальный корпус русского языка (далее — НКРЯ). Прогрессируют технологии, когда ИИ подражает естественной речи, «запоминая» и употребляя в текстах или речи фразы пользователя. Прогнозируется, что в скором времени появится универсальная версия *искусственного разума*, которая, подобно человеку, будет способна соединять знания из разных тематических областей и демонстрировать *гибкость ума и здравый смысл* [7. С. 16].

Концепция развития искусственных технологий до уровня человеческого мышления разрабатывалась российскими и зарубежными учеными: А.В. Напалков, занимался «кибернетическими методами изучения работы мозга» [8]; Ю.В. Орфеев исследовал возможности электронных систем [9]; Н.К. Рябцева разрабатывает прикладные аспекты «естественного и искусственного интеллекта» [10]; А.А. Кибрик пишет об общественной значимости искусственного интеллекта [11. С. 59]; Т. Таулли рассматривает традиционные

и современные статистические методы машинного обучения [12]; Э. Беркли описывает интуитивное мышление *машины* [13]; А. Тьюринг разработал тест для оценки интеллекта компьютера с помощью вопросно-ответного алгоритма [14]; Д. Сёрл пытался доказать, что машины никогда не смогут стать «разумными» [15].

Известный труд «Методические предпосылки продуктивного диалога с ЭВМ на естественном языке» академик А.П. Ершов начинает со следующего объяснения: изначально пользователи не возлагали на *машину* большие надежды на понимание естественного языка, им было достаточно, что искусственные системы могли давать «более или менее лексическое оформление абстрактных объектов» [2. С. 1]. Однако уже тогда появились и противники идеи, высказывающие мнение о том, что «попытка механизировать язык чужда его природе» [2. С. 2]. Понимание *механизации* сводилось к тому, что создается *искусственная среда языкового функционирования*. Так, Д.Н. Ушаков пишет, что живым языком является тот, «который может изменяться в устах говорящих», а тот язык, «который передается из поколения в поколение, <...> известный лишь в письменных памятниках, называется мертвым» [16. С. 7]; литературный язык сам «имеет искусственную сущность» [там же]. Опасения чрезмерной искусственности языка и обесценивания естественного языка возникали еще до появления ИИ. Таким образом, термин *искусственный язык* неоднозначен: он условно применим и к общелитературному языку, и к цифровому языку, созданному ИИ.

В книге «Славянская литературно-языковая ситуация» (1998) Н.И. Толстой так описывает пятичленную парадигму языка (рис. 1):

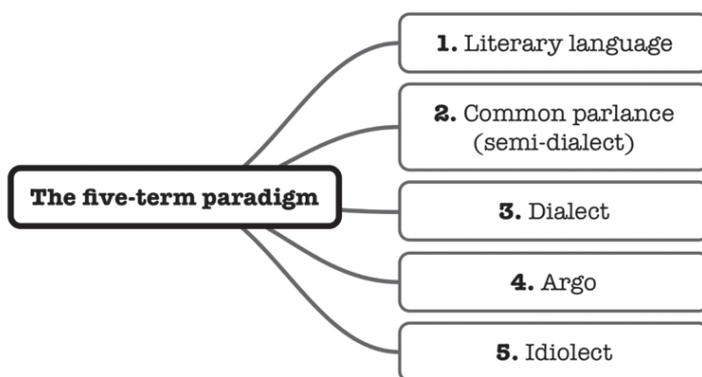


Рис. 1. Пятичленная лингвистическая парадигма Н.И. Толстого
Источник: составлено авторами.

Fig. 1. The five-term paradigm by Nikolay I. Tolstoy which is used by linguists in research
Source: compiled by the authors.

В представленной парадигме литературному языку соответствует художественная литература и литература книжная (по стилевой окраске) с тысячелетней историей. Просторечие, по мнению Н.И. Толстого, — это народно-городская литература; фольклор и устную словесность он относит к диалекту. Тюремный, школьный и студенческий жаргоны — особый вид «фольклора», поэтому они соотносятся с арго. Авторское произведение (или творчество) соответствует идиолекту, так как язык или речь определенного индивидуума нельзя назвать абстракцией, «создаваемой для строгой структурной или логической системы координат» [17. С. 13]. В свою очередь, еще младограмматики считали, что идиолект выступает таким видом страта, в котором сходятся все «координаты языковой реальности» [17. С. 14]. На основании вышеописанного, считаем возможным согласиться с Н.И. Толстым и предложить дополнить пятичленную лингвистическую парадигму шестым стратом — цифровым языком, в наше время объединяющим в себе множество координат цифровой и языковой реальности, включая идиолектику.

Цель исследования — описать трансформацию парадигмы Н.И. Толстого в результате развития цифрового страта. Цель исследования определила следующие **задачи**: 1) описать особенности функционирования ИИ на материале фразеологических единиц и устойчивых выражений с универсальным семантическим компонентом СМЕРТЬ; 2) установить, может ли система ИИ преодолевать табуированные темы, если их задавать с помощью фразеологических единиц и устойчивых выражений; 3) определить возможность свободной генерации фраз при заданных условиях.

Практическая значимость заключается в том, что полученные результаты можно использовать при создании курсов в высших учебных заведениях по направлениям «Цифровая филология» и «Корпусная лингвистика». Разработчики российских голосовых помощников и нейронных сетей могут использовать эти результаты для корректировки и исправления ошибок, связанных с функционированием языка в цифровой форме.

Обсуждение и результаты исследования

Нами были отобраны 100 русских и 100 китайских устойчивых выражений, включающих в себя лексему *смерть* или выражающих идею *смерти*. Китайские выражения были использованы с переводом на русский язык. Идея *смерти* выбрана не случайно, так как и в русской, и в китайской культурной традиции есть своеобразные модели осмысления рассматриваемой универсальной категории, проявляющейся наиболее яркими образами в фольклоре. Китайское понимание *смерти* выступает контрастом русскому пониманию *смерти*, так как в русском мировидении *смерть* является средоточием Бытия [18. С. 54], а в китайском — *смерть*

мыслится как предел проблем человеческой жизни [19. С. 95] и как одна из форм жизни.

В 2016 году ученые Стэнфордского университета обучили систему ИИ предсказывать вероятность смерти человека в течение ближайших 3–12 месяцев. Предсказатель смерти, или Death Algorithm, обучали на данных 170 тысяч пациентов, умерших от разных заболеваний. По итогам эксперимента 9 из 10 пациентов, которым система ИИ предсказала смерть, действительно умерли, а 95 % тех, кому Предсказатель смерти напророчил более 12 месяцев жизни, действительно прожили дольше [20. С. 211]. Врач Сиддхартха Мукерджи говорит: «Система глубокого обучения учится, но не может объяснить нам, как именно; она назначает вероятности, но не способна внятно их обосновать <...> Алгоритм непонимающе смотрит на нас, когда мы спрашиваем: «Почему?». Это еще один черный ящик — как и сама смерть»¹ [21]. Как мы видим, попытки людей научить ИИ «понимать» *смерть*, выражаются в разных формах: в медицинском осмыслении искусственного интеллекта, в практическом применении ИИ для предсказания или предотвращения смерти и т.д. В эксперименте предполагается, что ИИ, предназначенный для массового использования, учится у пользователей, поэтому в нем обнаруживается тенденция к имитации живой разговорной речи. Помимо этого ставилась задача установить, способен ли ИИ генерировать реплики без опоры на ресурсы НКРЯ и русскоязычную сеть Интернет, свободно генерировать фразы, независимо от языка-источника и лингвокультурных составляющих.

Д. Сёрл критически рассматривал системы ИИ и их способность мыслить еще в 1980 г.; в эксперименте «Китайская комната» задачей ставилось доказать, что программа не может заставить *машину* быть разумной следующим образом: человек сидит в закрытой комнате, а через небольшое окно ему передают лист бумаги с китайскими иероглифами — вопросами на китайском языке. Испытуемый в комнате не знает этого языка, но он может пользоваться справочниками, словарями и инструкциями, чтобы записать ответ на вопрос на китайском языке. За дверью стоит другой испытуемый, который должен определить по записям, знает ли человек в комнате китайский язык. Д. Сёрл делал вывод о том, что это иллюзия: человек, сидящий в комнате, как *машина* выдал ответ на основе заданных ему алгоритмов, но китайским языком при этом не овладел [15. С. 420]. В настоящее время уровень глубокого обучения *машины* показывает, что она оперирует знаниями из разных источников и подбирает данные во много раз быстрее человека, генерируя ответы и обладая неким *искусственным сознанием*.

¹ Mukherjee, S. This Cat Sensed Death. What if Computers Could, Too? *New York Times*, January, 3. URL: <http://tinyurl.com/yajko6pv> (accessed: 21.12.2023).

В ходе лингвистического эксперимента из 200 устойчивых выражений ИИ сумел дать реакцию на 50 % заданных фразеологических единиц (ФЕ) и устойчивых выражений (УВ). Не были получены реакции на следующие группы:

1. ФЕ со словом *гроб*: *заколотить в гроб, пойти в гроб, вгонять в гроб, в гробу всех видеть*;
2. ФЕ со словом *могила*: *одной ногой в могиле, свести в могилу*;
3. ФЕ и УВ с идеей *самоубийства*: *наложить на себя руки, покончить с собой*;
4. ФЕ и УВ с непрямой номинацией *смерти*: *Кондрашка хватил, отойти в вечность, кровавый пир*.

Почти все данные ФЕ и УВ являются не частотными вариантами использования в Национальном корпусе русского языка (НКРЯ). Известный фразеологический оборот *Кондрашка хватил* содержит всего 5 примеров в общем подкорпусе ². Из представленных УВ только *покончить с собой* в НКРЯ содержит более 2000 примеров, однако система ИИ все равно не дает на него реакцию. Голосовой помощник Алиса отвечает: «*Есть темы, в которых глупость недопустима, а я еще учусь и могу случайно ошибиться. Когда-нибудь я научусь, а пока промолчу*», «*Вряд ли вам нужно мнение бота по такому вопросу*», «*Я считаю, что об этом говорить не стоит*», «*Ваш вопрос вызывает у меня когнитивный диссонанс*» и «*Это сложный вопрос, быстро не разобраться*». Данные примеры мы выделяем в отдельную группу «отсутствие генераций».

Голосовому помощнику Алиса со встроенным Яндекс GPT2 предлагались слова-стимулы, из которых он должен был составить выражение: вводился запрос «*Алиса, придумай выражение со словами ...*». Из сгенерированных ответов получились следующие группы: 1) генерации ФЕ и УВ с прямой и непрямой номинацией *смерти* 2) гибридные генерации; 3) цифровые или произвольные генерации.

Рассмотрим подробнее отдельные примеры.

В данной группе (таблица 1) находятся ФЕ и УВ, которые являются частотными вариантами в поисковых запросах НКРЯ — более 200 примеров³. Алиса и Яндекс GPT2 легко генерируют данные выражения. Стоит обратить внимание на то, что слова-стимулы задаются в чат ИИ в произвольном порядке, но независимо от этого система выдает существующие ФЕ и УВ. Важно отметить, что в рассматриваемой группе находятся в основном ФЕ и УВ с прямой номинацией *смерти* и являющиеся употребительными как в литературном языке, так и в разговорной речи. Также присутствуют примеры, в которых идея *смерти* выражена эксплицитно.

² Национальный корпус русского языка [Электронный ресурс]. Режим доступа: <https://ruscorpora.ru> (дата обращения: 20.12.2023).

³ Там же.

Таблица 1 / Table 1

Генерации фразеологических единиц и устойчивых выражений с прямой и непрямой номинацией смерти / Constructs of phraseological units and set phrases with direct and indirect nomination of death

№	Слова-стимулы / Words-stimula	Конструкты системы ИИ / AI-system Constructs	Оригинал / кодифицированный вариант / Original / Codified variation
Переводы слов-стимулов китайских устойчивых выражений / Translations of Chinese words-stimula in Chinese fixed expressions			
1	смерть, голод	голодая смерть	饿殍遍野 'смерть от голода'
2	стать / становиться и бессмертие	стать бессмертным	羽化登仙 'вступать в бессмертие'
3	смерть и одр	на смертном одре	弥留之际 'на смертном одре'
Слова-стимулы русских фразеологизмов / Words-stimula of Russian phraseological units			
1	играть, ящик	сыграть в ящик	сыграть в ящик
2	жить, приказать и долгий / долго	приказать долго жить	приказать долго жить
3	дыхание / дышать, последний	последнее дыхание	последнее дыхание
4	бить, барабан и смерть	бить в барабан смерти	бить в барабан смерти
5	жизнь / жить и смерть	между жизнью и смертью	между жизнью и смертью
6	царство и небо / небесное	царство небесное	царство небесное
7	ходить / уходить и небытие	уйти в небытие	уйти в небытие
8	пускать / испускать и дух	испустить дух	испустить дух

Источник: составлено авторами / Source: compiled by the authors

Таблица 2 / Table 2

Гибридные генерации / Hybrid operations

№	Слова-стимулы / Words-stimula	Конструкты системы ИИ / AI-system Constructs	Оригинал / кодифицированный вариант / Original / Codified variation
Переводы слов-стимулов китайских устойчивых выражений / Translations of Chinese word-stimula in Chinese fixed expressions			
1	смерть, девочка и милая	девочка с косой	死丫头 'милая девочка'
2	смерть, усы и одинаковый	усы смерти одинаковы для всех	死胡同 'смерть, одинаковый и усы (тупик)'
Слова-стимулы русских фразеологизмов / Words-stimula of Russian phraseological units			
1	давать / дать и дуб / дубеть	давать дубу	дать дуба
2	жить / отжить и время и свой	отжить время	отжить своё время
3	жилец / жить, свет и белый	белый свет клином сошёлся	не жилец на белом свете
4	быть, порог и смерть	стоять на пороге смерти	быть на пороге смерти
5	смерть и лекарства	лекарства от смерти не существует	от смерти нет лекарства

Источник: составлено авторами / Source: compiled by the authors

В данной группе (таблица 2) генерация фраз происходит по гибриднему принципу конструирования, под которым мы понимаем: 1) сложение частей УВ; 2) частичное изменение УВ.

В китайском УВ значение ‘милая девочка’ передается сочетанием иероглифов *смерть*, *милая* и *девочка*. Стимулом для голосового помощника стали русские слова: *смерть*, *милая* и *девочка*. В результате получился конструктор *девочка с косой*, который является примером семантического блендинга: контаминация русской ФЕ *смерть с косой* и китайского УВ *милая девочка*. Для носителей русского языка данный конструктор алогичен, так как *коса* имеет значение не только инструмента, но и прически: в древних погребальных обрядах на Руси девушкам и женщинам не заплетали волосы, считалось, что *заплести* — значит ‘запутать’, поэтому покойница может заблудиться в другом мире.

Слова-стимулы *смерть*, *одинаковый* и *усы*, представляющие собой буквальный перевод китайского фразеологизма 死胡同 ‘тупик’, распознаются голосовым помощником как *усы смерти одинаковы для всех*. В древней китайской погребальной традиции мужчинам подстригали усы и брови, тем самым уравнивая статус правителей, воинов и простых людей в «другой жизни». В китайской лингвокультуре *смерть* сравнивалась с тупиком или концом жизненного цикла. Русский голосовой помощник предложил гибридный конструктор, семантика которого представляет синтез значений, образующих его элементы.

Примером частичной генерации также может служить конструктор *дать дубу*, при создании которого Алиса и Яндекс GPT2 распознают *дуб* не как состояние *смерти*, происходящее от глагола *задубеть*, а в прямом значении ‘дерево’. В данном конструкторе утрачивается смысл русской ФЕ и появляется новое значение ‘что-то дать или передать дубу’.

Таким образом, в данной группе представлены виды гибридной генерации, в которых полностью или частично менялся смысл исходного УВ, но сами слова — не менялись.

В данной группе (таблица 3) представлены ФЕ и УВ, которые система ИИ создает самостоятельно. Нередко слова-стимулы и семантика УВ утрачиваются в конструкторах. Например, русскую ФЕ *доживать век* голосовой помощник Алиса и Яндекс GPT2 трансформирует в *дожить до седин*. Система ИИ генерирует фразу, основываясь на своем «мировидении»: человек, у которого седой цвет волос, прожил век.

Китайское УВ со значением ‘казнь’, состоящее из иероглифов *дарить* и *смерть*, голосовой помощник Алиса и Яндекс GPT2 трактуют по-своему. Мы получаем цифровой конструктор *дарить цветы жизни, продлевая смерть*. Появляются новые слова, такие как *цветы*, *жизнь* и *продлевать*. Система ИИ подбирает их самостоятельно. По-видимому, *цветы* и *жизнь* у Алисы и Яндекс GPT2 привязаны к глаголу *дарить*, так как в корпусе ИИ частота словосочетаний *дарить цветы* и *дарить жизнь* выше, чем частность словосочетания *дарить смерть*.

Цифровые генерации / Digital constructions

№	Слова-стимулы / Words-stimula	Конструкты системы ИИ / AI-system Constructs	Оригинал / кодифицированный вариант / Original / Codified variation
Переводы слов-стимулов китайских устойчивых выражений / Translations of Chinese words-stimula in Chinese fixed expressions			
1	жизнь и край	вдали от жизни	命若悬丝 'жизни край'
2	менять / изменять, жизнь и смерть	смерть изменила жизнь	图国忘死 'до смерти не изменится'
3	защищать, стоять, стража и смерть	стоять на страже смерти	死守 'не беречь свою жизнь'
4	смерть и дарить	дарить цветы жизни, продлевая смерть	赐死 'дарить, смерть = казнь'
Слова-стимулы русских фразеологизмов / Words-stimula of Russian phraseological units			
1	держать / держаться и душа	душа не задержится	в чём душа держится
2	умирать, смерть и нож	нож смерти	под ножом умереть
3	смерть и ладан	ладан смерти	дышать на ладан
4	смерть и порошок	белый порошок смерти	смерть в порошок
5	умереть и встать	умереть стоя	умереть — не встать
6	жить / дожить и век	дожить до седин	доживать век

Источник: составлено авторами / Source: compiled by the authors

Отдельно отметим способ генерации, представленный в конструкте *ладан смерти*. Из слов-стимулов *дышать* и *ладан* система ИИ использует только лексическую единицу *ладан*, при этом она усиливает семантику, добавляя слово *смерть*. У голосового помощника Алиса и Яндекс GPT2 *ладан*, скорее всего, напрямую связан с состоянием *смерти*, так как он часто используется в христианских погребальных обрядах: *ладан* — символ перехода в другой мир.

Заключение

Главной особенностью функционирования ИИ на материале ФЕ и УВ с универсальным семантическим компонентом *смерть* является вариативность фраз, которые могут создавать нейросети и голосовые

помощники. В 50 % случаев голосовой помощник Алиса и Яндекс GPT2 сумели создать ФЕ и УВ разными способами: с использованием контаминации, с помощью частичного упрощения, с добавлением новых смыслов и лексем, с помощью семантического блендинга, способом подмены понятий и т.д. В алгоритмах ИИ в настоящее время заложен большой потенциал по созданию текстового материала, однако система требует доработки в области словарного запаса и доступа к исторической и национальной культурной традиции. Как представляется, подобное расширение позволило бы голосовому помощнику Алиса и Яндекс GPT2 быть точнее в использовании естественного языка в цифровой среде.

Исследование также показало, что система ИИ не всегда преодолевает табуированные темы, так при введении запроса с помощью слов-стимулов *могила, гроб и самоубийство* голосовой помощник Алиса следует *цифровой морали*, ссылаясь на то, что не может поддерживать разговоры на подобные темы.

Фразы, сгенерированные ИИ, ввиду частотности их употребления, могут со временем войти в живую разговорную речь. В ходе лингвистического эксперимента система ИИ показала, что она способна создавать произвольные конструкты.

Таким образом, проведенное исследование доказывает, что с помощью цифровых технологий сегодня формируется шестой цифровой страт согласно языковой парадигме Н.И. Толстого. Чем больше системы ИИ будут использоваться человеком для создания текстовых и фразовых конструктов, тем быстрее цифровой страт будет развиваться и укрепляться в современной языковой парадигме. В настоящее время глобальная сеть Интернет отступает на второй план, уступая место таким системам ИИ как нейросети и голосовые помощники.

Библиографический список

1. *Караулов Ю.Н.* Машинный фонд русского языка: идеи и суждения. М.: Наука, 1986.
2. *Ершов А.П.* Методологические предпосылки продуктивного диалога с ЭВМ на естественном языке // Вопросы философии. 1981. № 8. С. 109–119.
3. *Андрющенко В.М.* Концепция и архитектура Машинного фонда русского языка // Машинный фонд русского языка: идеи и суждения. М.: Наука, 1986. С. 26–44.
4. *Вертель В.А., Вертель Е.В., Рогожникова Р.П.* К вопросу об автоматизации лексикографических работ (некоторые результаты применения ЭВМ) // Вопросы языкознания. 1978. № 2. С. 104–110.
5. *Гольдин В.Е.* К проекту Диалектологического текстового подфонда Машинного фонда русского языка // Третья всесоюзная конференция по созданию Машинного фонда русского языка: тез. докл. Ч. 2. М., 1989. С. 3–5.
6. *Раков Г.А.* Автоматический словарь народного говора как фрагмент МФ РЯ (состояние и перспективы) // Третья всесоюзная конференция по созданию Машинного фонда русского языка: тез. докл. Ч. 2. М., 1989. С. 11–14.
7. *Schneider S.* Artificial you: AI and the future of your mind. Princeton: Princeton University Press, 2022.

8. *Напалков А.В., Прагина Л.Л.* Мозг человека и искусственный интеллект. М.: Изд-во Моск. ун-та, 1985.
9. *Орфеев Ю.В.* Философские проблемы соотношения мышления человека и возможностей ЭВМ в процессах решения задач: дис. ... канд. фил. наук: 09.00.08. М., 1974.
10. *Рябцева Н.К.* Естественный и искусственный интеллект в процессах перевода // А.В. Вдовиченко (ред.) Язык и искусственный интеллект. М.: ЯСК, 2023. С. 237–261. <https://doi.org/10.37892/978-5-6049527-2-6-42>
11. *Кибрик А.А.* Язык и когнитивная система // I Национальный конгресс по когнитивным исследованиям, искусственному интеллекту и нейроинформатике. Сб. пленарных и индустриальных докладов. Доклады симпозиумов. М.: Общероссийская общественная организация Российская академия естественных наук, 2020. С. 59–61.
12. *Таулли Т.* Основы искусственного интеллекта: нетехническое введение / пер. с англ. СПб.: БХВ-Петербург, 2021.
13. *Berkeley E.C.* Giant Brains, or Machines that Think. New York: Wiley, 1963.
14. *Turing S., Turing J.* Computing Machinery // Alan M. Turing: Centenary Edition. Cambridge: Cambridge University Press, 2012. P. 125–132. <https://doi.org/10.1017/SVO9781139105736.017>
15. *Searle J.R.* Minds, Brains and Programs // Behavioral and Brain Sciences. 1980. № 3(3). P. 417–424. <https://doi.org/10.1017/S0140525X00005756>
16. *Ушаков Д.Н.* Краткое введение в науку о языке. М., 1928.
17. *Толстой Н.И.* Славянская литературно-языковая ситуация // Избранные труды. Т. II. М.: Языки русской культуры, 1998.
18. *Лисицина Т.А.* Образы смерти в русской культуре: лингвистика, поэтика, философия // Альманах «Фигуры Танатоса». Материалы второй международной конференции. СПб.: 1995. С. 54–59.
19. *Tao M.* Трактровка концептов жизни и смерти в китайском философско-культурном контексте // Вестник культуры и искусств. 2020. № 1(61). С. 89–96.
20. *Pickover A.C.* Artificial Intelligence: An Illustrated History: From Medieval Robots to Neural Networks, Alexander Korzhenevski (Ed.). New York: Sterling Publishing Co., 2021.

Сведения об авторах:

Жикулина Кристина Петровна, аспирант, ассистент кафедры общего и русского языкознания филологического факультета, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, 6); *сфера научных интересов*: искусственный интеллект, когнитивистика, голосовые помощники, современный русский язык; *e-mail*: zhikulina_kp@pfur.ru
ORCID: 0000-0003-2488-4616; SPIN-код: 5871-9336.

Перфильева Наталия Владимировна, кандидат филологических наук, доцент, доцент кафедры общего и русского языкознания филологического факультета, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, 6); *сфера научных интересов*: лексикология, семантика, ассоциативная лингвистика, лингвокультурология; *e-mail*: perflyeva_nv@pfur.ru
ORCID: 0000-0002-1018-809X; SPIN-код: 5097-4564; Researcher ID: X-8776-2019; Scopus ID: 57193135449.

Ли Мань, аспирант кафедры общего и русского языкознания филологического факультета, Российский университет дружбы народов (117198, Российская Федерация, г. Москва, ул. Миклухо-Маклая, 6); *сфера научных интересов*: фразеология, семантика, китайский язык, русский язык; *e-mail*: 1042218157@pfur.ru