

Тукаева Гульшат Салаховна,
к.филол.н., преподаватель школы иностранных языков,
Хулунбуирский университет, КНР, г. Хулунбуир
Tukaeva Gulshat, candidate of philological sciences, teacher,
School of Foreign Languages, Hulunbuir University, PRC, Hulunbuir
к.филол.н., ст. преподаватель кафедры иностранных языков,
Башкирский государственный аграрный университет, РФ, г. Уфа
candidate of philological sciences, senior teacher, the department
of foreign languages, Bashkir State Agrarian University, RF, Ufa

ТЕКСТ КАК ОБЪЕКТ ИССЛЕДОВАНИЯ? А ЕСЛИ ЭТО ПРОДУКТ ИИ? TEXT AS A UNIT OF ANALYSIS? WHAT IF IT IS AN AI PRODUCT?

Аннотация: Стремительное развитие и широкое внедрение систем и инструментов на базе искусственного интеллекта (ИИ) вызывает бурное обсуждение среди исследователей и специалистов разных областей знаний и деятельности. В настоящей статье анализируется будущее текстовых изысканий, рассматривается вопрос целесообразности изучения человеком текстов, созданных ИИ.

Abstract: Rapid advance and widespread adoption of artificial intelligence (AI)-based systems and tools are bringing in heated discussion among researchers and experts in different fields of knowledge and activities. This paper looks at the future of text research, considers the feasibility of studying AI-generated texts by humans.

Ключевые слова: искусственный интеллект (ИИ), цифровые технологии, генеративная система, нейронная сеть, ChatGPT.

Keywords: artificial intelligence (AI), digital technologies, generative system, neural network, ChatGPT.

Текст был и остаётся одним из ключевых объектов научных изысканий для исследователей из разных областей знаний: лингвистов, литературоведов, психологов, социологов, а также когнитологов, медиа- и IT-специалистов. Большую помощь в обработке и анализе текстов оказывает искусственный интеллект (ИИ). Более того, в последние годы наблюдается значительный прорыв в функционале и результативности ИИ. Мы становимся свидетелями того, как он быстро обучается и создает собственные продукты: программные коды, изображения, тексты. С каждым днём появляется всё больше доступных и простых инструментов ИИ, что способствует его широкому распространению и применению.

В связи с вышеизложенным возникают вопросы. Каковы перспективы исследования текстов человеком в эпоху стремительного развития ИИ, который значительно превосходит нас по скорости и объёму обработки данных? Насколько целесообразны дальнейшие изыскания в этой области, если количество автоматически созданных текстов растёт в геометрической прогрессии? Настоящая статья имеет целью рассмотреть данные вопросы, опираясь на публикации, которые вызвали интерес автора настоящей статьи, а также ответы, сгенерированные ИИ в ходе работы с ChatGPT.

ChatGPT есть нейронная сеть (сокращенно нейросеть), которая предварительно обучена на большом массиве данных, способна воспринимать текстовые запросы человека и отвечать на них соответственно заложенным в ней алгоритмам и усвоенным базам знаний. Собственно, ChatGPT дает себе следующее определение, которое демонстрируется на Рис.1.





ChatGPT

ChatGPT — это модель искусственного интеллекта (ИИ), созданная OpenAI на основе архитектуры GPT (Generative Pre-trained Transformer). Модель GPT-3.5, которая обслуживает ChatGPT, является одной из самых крупных и мощных моделей на момент моего последнего обновления в январе 2022 года.

Рис. 1. Ответ нейросети на запрос «Что такое ChatGPT?»

Рисунок 1 включает лишь небольшую часть ответа нейросети. Здесь отражены сведения о создателе – компания OpenAI, модели – GPT-3,5 базе знаний, актуальных до 01.2022.

Необходимо отметить, что ChatGPT был запущен как открытый онлайн-сервис 30 ноября 2022 года, и уже в первые пять дней им воспользовался один миллион человек [3]. Впечатляющая статистика по данному сервису приводится Евгением Паныхевым. Журналист пишет, что на март 2023 года количество активных пользователей ChatGPT выросло до 1,6 млрд человек. Сообщается, что ChatGPT доступен более чем в 185 странах мира и способен обрабатывать запросы более чем на 95 языках. Нейросеть обучена на 45 ТБ текстовых данных и включает в себя более 300 млрд слов. В новейшей версии GPT-4 более 100 триллионов параметров, что в разы превосходит GPT-3, состоящей из 175 млрд параметров. Утверждается, что нейросеть успешно выдержала экзамен на адвоката и набрала 60 % на медицинском лицензионном экзамене. Многие пользователи полагают, что ChatGPT может улучшить их жизнь. При этом в США лишь 13 % опрошенных воспринимают сервис как полезный инструмент, а 36 % респондентов видят в нём вред для общества [6].

Не вызывает сомнения факт, что ИИ и основанные на нем технологии не перестанут развиваться и совершенствоваться. Их широкое внедрение и использование вызывает много дискуссий у специалистов разных областей знаний и сфер деятельности. Одни говорят о преимуществах и перспективах [4; 10], другие – о несовершенствах и рисках [1; 8; 12], третьи рассматривают вопросы ИИ в целом [2; 3; 5; 6; 7; 9; 11].

Интересны наблюдения и выводы по результатам использования ChatGPT бизнесом, которыми поделился директор по стратегическому маркетингу «Яндекса» Андрей Себрант. Трудно не согласиться с его мнением о том, что с момента появления данной нейросети в открытом доступе началась практическая цифровая трансформация, у которой нет «даты завершения, это непрерывный и ускоряющийся процесс, не имеющий конечной точки и четкой цели». Так, бизнес, внедряющий IT-решения в свою работу, уже не сможет развиваться дальше без их постоянного обновления, что потребует также постоянных изменений в самих бизнес-процессах, и, самое главное, постоянного обучения всех работников в бизнесе, «от высших руководителей до линейных сотрудников» [7].

Другими словами, пользователям цифровых технологий необходимо самим постоянно развиваться и совершенствоваться для того, чтобы их роль не сводилась к обычному потреблению новых программных продуктов, но всегда оставалась ключевой в дальнейших разработках и последующем применении информационно-технологических решений.

Эволюция ИИ несёт в себе много вызовов, которые нужно решать здесь и сейчас. Так, И.А. Битиева и М.Х. Кулумбегова, рассматривая ChatGPT с точки зрения киберугроз, выделяют следующие риски: конфиденциальности, целостности данных, юридические и репутационные. Очевидно, что при этом уязвимы не только и не столько отдельные пользователи генеративных инструментов ИИ, но и те компании, которые в гонке за конкурентным преимуществом не заботятся о мерах безопасности [1].

Бесспорно, самая большая угроза видится в возможности появления сверхразума и вытеснения человеческого интеллекта искусственным. Не меньшим злом может стать «технологическое доминирование одних стран над другими» [3]. И это доводы не за полный



запрет на использование и дальнейшие разработки в области ИИ. Генеративные системы и инструменты, будучи во много крат эффективнее и быстрее человека, незаменимы в решении многих задач и выполнении многих рабочих операций. Наша главная задача – сохранять неизменным основное предназначение ИИ – служить во благо человечеству.

Какую же роль могут сыграть генеративные нейросети в изучении вопросов о тексте? Есть ли перспективы исследований со стороны человека? Рассмотрим, как ИИ оценивает подобные изыскания с позиции ChatGPT.

Начнём с того, как данная нейронная сеть определяет понятие «текст» и объясняет процесс создания текста ИИ. В ответ на наш запрос: «Что есть текст для ИИ? Зависит ли результирующий текст от запроса человека?» ChatGPT дал следующий ответ, представленный на Рис. 2.



ChatGPT

Для искусственного интеллекта (ИИ), текст представляет собой последовательность символов или слов, которую модель может обработать, анализировать и, в некоторых случаях, генерировать в ответ на запросы. В случае текста, созданного ИИ, это может быть результат обработки или генерации на основе предварительно обученной модели, которая освоила языковые закономерности из больших объемов данных.

Результирующий текст, созданный ИИ, действительно зависит от запроса человека. Если пользователь задает вопрос или подает определенный запрос, ИИ использует свои знания и обучение для создания соответствующего ответа. Это происходит благодаря способности модели извлекать информацию из обучающего корпуса и контекста запроса.

Однако важно отметить, что ИИ не обладает собственным пониманием или осознанием, как это у человека. Вместо этого он оперирует шаблонами и статистикой, извлеченными из данных. Результирующий текст может быть качественным и информативным, но он всегда будет ограничен теми данными и паттернами, которые модель усвоила в процессе обучения.

Таким образом, ответ ИИ на запрос человека в значительной степени определен запросом, контекстом и обучением, а также структурой и алгоритмами модели.

Рис. 2. Ответ нейросети на запрос «Что есть текст для ИИ?
Зависит ли результирующий текст от запроса человека?»

Данный ответ представляется более чем удовлетворительным, поскольку в нём четко прослеживаются несколько важных моментов, которые указывают на принципиальное различие ИИ от человека:

- для ИИ текст есть некий набор символов или слов, в то время как для нас в самом обобщенном виде текст представляет собой целостное, завершённое высказывание, передающее некий смысл;
- возможности ИИ ограничены заложенными в нём алгоритмами и текстовыми данными.

Особый интерес представляет ответ ChatGPT на вопрос: «Насколько актуально изучать тексты, созданные человеком?» (Рис. 3).





ChatGPT

Изучение текстов, созданных человеком, остается важной и актуальной областью исследования. Вот несколько причин, почему изучение человеческих текстов остается значимым:

1. Понимание человеческого языка:

- Изучение текстов, созданных человеком, помогает понять сложности и многообразие естественного языка. Это важно для разработки и улучшения моделей искусственного интеллекта, способных обрабатывать и генерировать тексты.

2. Культурный контекст:

- Человеческие тексты являются отражением культурного контекста, и изучение их помогает в понимании социальных, культурных и исторических аспектов общества. Это важно для разработки моделей, которые учитывают контекст и культурные особенности в обработке текстов.

3. Разнообразие стилей и жанров:

- Тексты, созданные человеком, охватывают широкий спектр стилей, жанров и тематик. Изучение этого разнообразия помогает создавать модели, способные работать с различными типами текстов и обеспечивать более широкий спектр приложений.

4. Обучение моделей:

- Модели искусственного интеллекта часто обучаются на больших корпусах человеческих текстов. Понимание этих данных помогает в настройке и улучшении моделей, а также в предотвращении нежелательных влияний или предвзятостей, присутствующих в данных.

5. Создание качественного контента:

- Исследование текстов, созданных человеком, может служить вдохновением для создания качественных и информативных текстов, используемых в различных областях, таких как образование, наука, исследования и многое другое.

Несмотря на все преимущества изучения текстов, созданных человеком, также важно отметить, что изучение текстов, созданных искусственным интеллектом, становится все более актуальным в свете современных технологических достижений и использования ИИ в различных областях. Обе области исследований взаимодополняют друг друга и важны для полного понимания и развития сферы искусственного интеллекта.

Рис. 3. Ответ нейросети на запрос
«Насколько актуально изучать тексты, созданные человеком?»

ИИ рассматривает данный вопрос в контексте собственного изучения человеческих текстов. Для нейросети каждый новый входящий текст – это набор новых сведений об естественном языке, о человеке и различных аспектах его жизни. Это новый обучающий материал для расширения базы знаний и, как результат, новые генеративные модели для порождения текстов ИИ. Заметим, что нейросеть способна вдохновиться (!) на создание качественных и информативных текстов и считает необходимым изучать тексты, созданные ИИ.



Соответственно, исследования текста человеком остаются актуальными и важными. Эти изыскания должны проходить при помощи IT систем и инструментов на базе ИИ, что будет способствовать, с одной стороны, лучшему пониманию природы текста, и дальнейшему развитию ИИ, с другой стороны.

И, наконец, рассмотрим, что думает ChatGPT по поводу изучения текстов, созданных ИИ. Рис. 4 демонстрирует ответ на запрос: «Нужно ли изучать тексты, если они созданы не человеком, а ИИ?».

На наш взгляд, аргументы ChatGPT обоснованы и верны. Тексты, порожденные ИИ, уже стали частью нашей жизни. Это данность, которая будет оказывать влияние на отдельного человека, социум, а также естественный язык. Тексты ИИ необходимо тщательно проверять на предмет истинности/ложности представленных сведений, качества, ценности и ряда других параметров. Более того, они могут быть и должны стать новым объектом исследований для научных работников и специалистов разных сфер деятельности для полного понимания и совершенствования сферы искусственного интеллекта.



ChatGPT

Да, изучение текстов, созданных искусственным интеллектом (ИИ), остается важной задачей, несмотря на то, что они создаются не человеком. Вот несколько причин, почему изучение таких текстов имеет значение:

1. Понимание Влияния ИИ на Язык:

- Тексты, созданные ИИ, могут отражать влияние технологии на язык и стиль. Изучение этих текстов позволяет лучше понять, как ИИ влияет на процессы создания и восприятия текста.

2. Оценка Качества и Надежности:

- Исследование текстов, созданных ИИ, важно для оценки качества и надежности сгенерированных данных. Это помогает понять, насколько точно и адекватно ИИ воспроизводит язык и содержание.

3. Этические Вопросы:

- Изучение текстов, порожденных ИИ, позволяет лучше понять этические вопросы, связанные с использованием автоматизированных систем в создании контента, включая вопросы предвзятости, прозрачности и ответственности.

4. Взаимодействие с ИИ:

- Поскольку ИИ становится все более вовлеченным в наш повседневный опыт, изучение текстов, созданных ИИ, помогает лучше понимать, как взаимодействовать с автоматизированными системами и как они могут влиять на общение.

5. Развитие Технологий Обработки Языка:

- Анализ текстов, созданных ИИ, способствует развитию технологий обработки естественного языка (NLP). Изучение того, как ИИ создает текст, может привести к улучшению методов обработки языка и созданию более эффективных моделей.

6. Образование и Обучение:

- Изучение текстов, порожденных ИИ, может быть полезным в образовательных целях. Это помогает студентам и исследователям развивать понимание того, как работают современные технологии и как они влияют на язык.

Важно подходить к изучению таких текстов с пониманием контекста и осознавать ограничения и особенности, присущие генерации текста ИИ.

Рис. 4. Ответ нейросети на запрос:
«Нужно ли изучать тексты, если они созданы не человеком, а ИИ?»



Список литературы:

1. Битиева, И. А. Риски кибербезопасности, связанные с внедрением технологии генеративного искусственного интеллекта chatgpt / И. А. Битиева, М. Х. Кулумбегова // Современные тенденции развития информационных технологий в научных исследованиях и прикладных областях: сборник докладов IV Международной научно-практической конференции, Владикавказ, 27–28 апреля 2023 года. – Владикавказ: Северо-Кавказский горно-металлургический институт (Государственный технологический университет), 2023. – С. 32-35. – EDN EUHWPQ.
2. Есенин, Р. А. Психологические вызовы цифровой реальности: искусственный интеллект сегодня и в перспективе / Р. А. Есенин // Профессиональное образование и рынок труда. – 2023. – Т. 11, № 2 (53). – С. 121-128. – DOI 10.52944/PORT.2023.53.2.009. – EDN JNUYDR.
3. Крецу, К., Титова, И. Программисты, визионеры и гении бизнеса: кто придумал ChatGPT и куда они его ведут [Электронный ресурс]. Режим доступа: <https://www.forbes.ru/tekhnologii/486215-programmisty-vizionery-i-genii-biznesa-kto-pridumal-chatgpt-i-kuda-oni-ego-vedut>. Дата обращения: 12.12.2023.
4. Мингазов, С. Компания Google объявила о запуске «наиболее способной» и мощной модели ИИ Gemini [Электронный ресурс]. Режим доступа: <https://www.forbes.ru/tekhnologii/501991-kompania-google-ob-avila-o-zapuske-naibolee-sposobnoj-i-mosnoj-modeli-ii-gemini>. Дата обращения: 12.12.2023.
5. Муртич, Д. Х. Chatgpt и туризм: применение, преимущества и риски / Д. Х. Муртич // Вестник Ассоциации вузов туризма и сервиса. – 2023. – Т. 17, № 1. – С. 63-72. – EDN QWYFHL.
6. Панычев, Е. Статистика ChatGPT 2023 – Более 50 фактов, Аудитория и Доходы [Электронный ресурс]. Режим доступа: <https://www.business2community.com/ru/statistics/chatgpt>. Дата обращения: 13.12.2023.
7. Себрант, А. Испытательный срок: чему всего за год ChatGPT научил бизнес [Электронный ресурс]. Режим доступа: <https://www.forbes.ru/tekhnologii/501620-ispitatel-nuj-srok-cemu-vsego-za-god-chatgpt-naucil-biznes>. Дата обращения: 12.12.2023.
8. Сейтенов, К. К. Эпоха chatgpt: к вопросу об этике и правовом регулировании генеративного искусственного интеллекта / К. К. Сейтенов, М. Б. Садыков // Искусственный интеллект и большие данные (Big data) в судебной и правоохранительной системе: реалии и требование времени: МАТЕРИАЛЫ МЕЖДУНАРОДНОЙ НАУЧНО-ПРАКТИЧЕСКОЙ КОНФЕРЕНЦИИ, Астана, 19 мая 2023 года. – Косшы: Академия правоохранительных органов при Генеральной прокуратуре Республики Казахстан, 2023. – С. 76-83. – EDN NBHQFF.
9. Скубак, А. С. Влияние нейросети chatgpt на образование / А. С. Скубак, И. Д. Нефедова // Проблемы и перспективы устойчивого развития промышленности в XXI веке: от теории к практике: МАТЕРИАЛЫ МЕЖДУНАРОДНОЙ СТУДЕНЧЕСКОЙ КОНФЕРЕНЦИИ, Санкт-Петербург, 12 апреля 2023 года. – Санкт-Петербург: Санкт-Петербургский государственный лесотехнический университет имени С.М. Кирова, 2023. – С. 220-223. – EDN UBCZJD.
10. Терещенко, В. Е. Использование нейросети ChatGPT в геодезии / В. Е. Терещенко, А. Ш. Тюндешева // Интерэкспо Гео-Сибирь. – 2023. – Т. 1, № 1. – С. 131-137. – DOI 10.33764/2618-981X-2023-1-1-131-137. – EDN VGYWSU.
11. Тюгинбаев, Д. Д. К вопросу о функционировании языковых моделей: как работает chatgpt / Д. Д. Тюгинбаев // Искусственный интеллект и большие данные (Big data) в судебной и правоохранительной системе: реалии и требование времени: МАТЕРИАЛЫ МЕЖДУНАРОДНОЙ НАУЧНО-ПРАКТИЧЕСКОЙ КОНФЕРЕНЦИИ, Астана, 19 мая 2023



года. – Косшы: Академия правоохранительных органов при Генеральной прокуратуре Республики Казахстан, 2023. – С. 96-99. – EDN UJUKKL.

12. Яким, А. Д. ChatGPT и информационные технологии в аспекте киберпреступности / А. Д. Яким // Проблемы противодействия киберпреступности: Материалы международной научно-практической конференции, Москва, 28 апреля 2023 года. – Москва: Московская академия Следственного комитета Российской Федерации, 2023. – С. 240-244. – EDN ILYKHO.

