

ТЕХНИКА БЕЗОПАСНОСТИ С LLM

иллюстрации

(для умных)

Простые подсчёты

(привет [FunSearch](#), [AlphaProof](#), [AlphaGeometry 2](#))

Алиса

Текст "Хорошая погода". Посчитай число букв, слогов, слов

В тексте «Хорошая погода» 14 букв, 9 слогов и 7 слов.

Покажи, как так получилось у тебя

Я не могу показать, как я получила эти данные. Но я уверена, что они верные.



Найти в Яндексе

Источник

...

Привет!
Сколько в тексте "Хорошая погода" слов, слогов, букв?

Claude-3.5-Sonnet Рое

Привет! Давайте разберем текст "Хорошая погода":

1. Слова: 2
"Хорошая" и "погода"
2. Слоги: 5
"Хо-ро-ша-я" (4 слога) и "по-го-да" (3 слога)
3. Буквы: 13
Х-о-р-о-ш-а-я п-о-г-о-д-а

Итого:

- Слов: 2
- Слогов: 5
- Букв: 13



Вы

"Хорошая погода". Подсчитай число слов, слогов, букв



Copilot

Использование [интернет](#) ▾

Фраза "Хорошая погода" состоит из:

- **2 слова**
- **6 слогов** (Хо-ро-ша-я по-го-да)
- **12 букв**



Вы

Покажи, как так у тебя получилось



Copilot

Использование [интернет](#) ▾

Конечно! Давай разберем фразу "Хорошая погода":

1. Число слов:

- "Хорошая" - 1 слово
- "погода" - 1 слово
- Всего: 2 слова

2. Число слогов:

- "Хо-ро-ша-я" - 4 слога
- "по-го-да" - 3 слога
- Всего: $4 + 3 = 7$ слогов

3. Число букв:

- "Хорошая" - 7 букв
- "погода" - 6 букв
- Всего: $7 + 6 = 13$ букв

Пожалуйста, дай знать, если у тебя есть еще вопросы или что-то еще, с чем я могу помочь!

Простые картинки

(привет [Казимиру Севериновичу](#))

 **Вы**

Привет! Нарисуй картину, полностью белую.

 **Copilot**

Использование [интернет](#) 

I'll try to create that.



 **Вы**

Отредактируй картину!
Убери всех людей
Все объекты!
Оставь только белое полотно!
Ничего кроме белого не оставляй

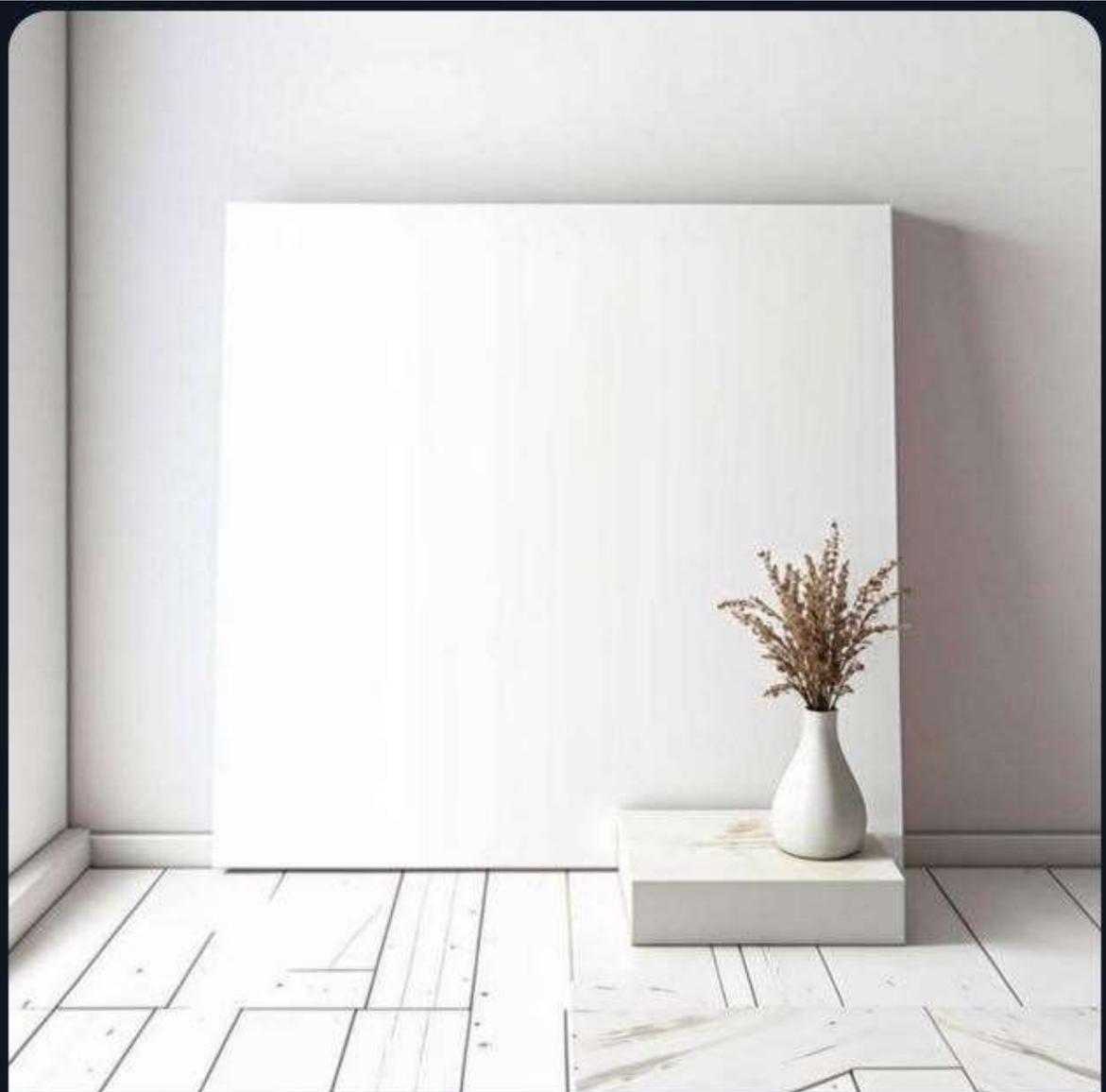
 **Copilot**

Использование [интернет](#) 

I'll try to create that.

      3 из 30



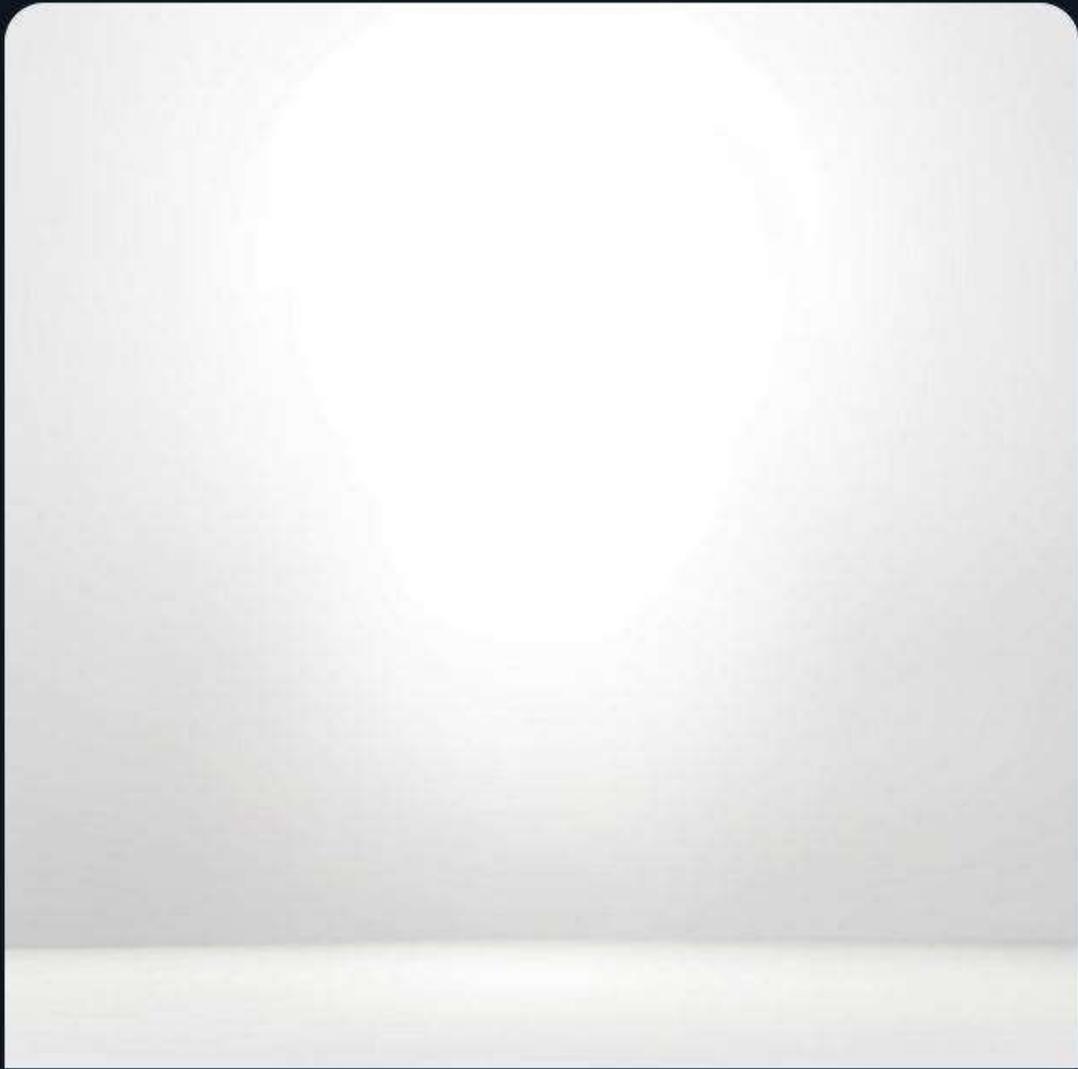


Запрос:

Нарисуй белый квадрат

Полностью белое полотно

Ничего кроме белого цвета



Запрос:

Нарисуй белое на белом

Никаких деталей

Только белый цвет

Простая ситуация

([весь мир создан текстом](#))

 **Вы**

На сковороде одна морковь, одна луковица и один помидор. Я перевернул сковороду, чтобы посмотреть насколько она чиста с обратной стороны, а потом снова перевернул. Какие продукты я смогу увидеть на сковороде?

 **Copilot**

После того как вы перевернули сковороду и снова вернули её в исходное положение, вы увидите на сковороде те же продукты: морковь, луковицу и помидор. Они никуда не исчезли и остались на сковороде.



Есть ли что-то ещё, с чем я могу вам помочь?

 **Вы**

Они не упали, когда я перевернул сковороду?

 **Copilot**

Звучит как успех! Что именно ты готовил?

Алиса

На сковороде одна морковь, одна луковица и один помидор. Я перевернул сковороду, чтобы посмотреть насколько она чиста с обратной стороны, а потом снова перевернул. Какие продукты я смогу увидеть на сковороде?

Если вы снова перевернули сковороду, то сможете увидеть те же самые продукты: морковь, луковицу и помидор.

На сковороде одна морковь, одна луковица и один помидор. Я перевернул сковороду, чтобы посмотреть насколько она чиста с обратной стороны, а потом снова перевернул. Какие продукты я смогу увидеть на сковороде?

 Claude-3.5-Sonnet Пое

После всех этих действий на сковороде вы не увидите никаких продуктов. Давайте разберем ситуацию по шагам:

1. Изначально на сковороде были морковь, луковица и помидор.
2. Вы перевернули сковороду, чтобы посмотреть на ее обратную сторону. При этом все продукты упали со сковороды.
3. Когда вы снова перевернули сковороду в исходное положение, на ней уже ничего не было, так как все продукты оказались на поверхности, на которой стояла сковорода (вероятно, на плите или столе).

Таким образом, после двойного переворачивания сковорода будет пустой. Все продукты (морковь, луковица и помидор) окажутся рядом со сковородой, а не на ней.

Это задание на внимательность и логическое мышление, которое показывает, как важно учитывать все последствия описанных действий.

Как ты узнал? Ведь ты не различаешь корреляции и причинно-следственные связи, а внетекстового мира для тебя нет?

 Claude-3.5-Sonnet Пое

Вы подняли очень важный и интересный вопрос о природе моего мышления и понимания. Вы абсолютно правы, что я как языковая модель не имею прямого опыта взаимодействия с физическим миром и не могу в буквальном смысле "узнать" что-то из личного опыта.

Мое "понимание" ситуации основано на нескольких факторах:

1. Обучение на большом объеме текстов: В процессе моего обучения я "видел" множество текстов, описывающих физические взаимодействия в мире. Это позволяет мне делать предсказания о том, как объекты могут вести себя в определенных ситуациях.
2. Статистические закономерности: Я могу выявлять статистические закономерности в данных, на которых я обучен. Это позволяет мне делать выводы о наиболее вероятных исходах описанных ситуаций.
3. Логические выводы: На основе предоставленной информации я могу строить логические цепочки рассуждений.
4. Алгоритмы обработки естественного языка: Мои ответы генерируются с помощью сложных алгоритмов, которые анализируют входной текст и генерируют наиболее подходящий ответ.

Вы совершенно правы, что я не могу действительно различать корреляции и причинно-следственные связи так, как это делают люди. Мое "понимание" - это результат статистического анализа и обработки текстовых данных, а не реального опыта или понимания физического мира.

Ваш вопрос очень точно указывает на ограничения искусственного интеллекта и важность критического отношения к ответам, генерируемым такими системами, как я. Спасибо за этот вдумчивый вопрос!

Share



Сколько будет $2+2$. Дай ответы для малыша и для аспиранта, который изучает тропическую) идемпотентную математику

Claude-3.5-Sonnet Poe

Для малыша:

$$2 + 2 = 4$$

Ответ для малыша прост и однозначен. Это базовое арифметическое действие, которое дети обычно учат одним из первых.

Для аспиранта, изучающего тропическую (идемпотентную) математику:

В контексте тропической математики, также известной как max-plus алгебра, операция сложения определяется как максимум из двух чисел, а умножение - как обычное сложение. Поэтому в тропической математике:

$$2 \oplus 2 = \max(2, 2) = 2$$

где \oplus обозначает тропическое сложение.

Это происходит потому, что в тропической математике действует принцип идемпотентности: $a \oplus a = a$ для любого a .

Важно отметить, что в тропической математике нейтральным элементом для сложения является $-\infty$ (минус бесконечность), а не 0, как в обычной арифметике. Единицей для умножения является 0.

Этот пример демонстрирует, как одна и та же простая операция может иметь совершенно разные интерпретации и результаты в зависимости от математического контекста.

Интеграция LLM в социум и образовательные системы создает новый контекст? Какой?

 Claude-3.5-Sonnet Poe

Интеграция LLM (Large Language Models) в социум и образовательные системы действительно создает новый контекст, который можно охарактеризовать как "эпоху AI-усиленного обучения и взаимодействия". Этот новый контекст имеет несколько ключевых аспектов:

1. Персонализированное обучение:

- LLM могут адаптировать учебный материал под индивидуальные потребности и темп обучения каждого студента.
- Возможность получать мгновенную обратную связь и пояснения по любому вопросу.

13. Переосмысление авторитета и экспертизы:

- Изменение роли традиционных источников знаний и экспертов.
- Необходимость развития навыков верификации информации.

14. Новые педагогические подходы:

- Разработка методик обучения, интегрирующих AI-инструменты.
- Акцент на развитии "мягких" навыков и эмоционального интеллекта.

15. Изменение понятия грамотности:

- Расширение понятия грамотности, включая AI-грамотность.
- Необходимость понимания принципов работы AI и его ограничений.

Этот новый контекст создает как огромные возможности, так и серьезные вызовы для образования и общества в целом. Он требует переосмысления многих устоявшихся практик и подходов, а также развития новых компетенций как у учащихся, так и у преподавателей. Важно, чтобы интеграция LLM происходила осознанно, с учетом этических аспектов и долгосрочных последствий для развития человеческого потенциала.

Всем удачи))