

# | Границы применимости LLM |

Ключевые особенности применения Chat GPT и других

LLM

Марк Андреев @ Conundrum AI  
mark.andreev@gmail.com

Май 2024

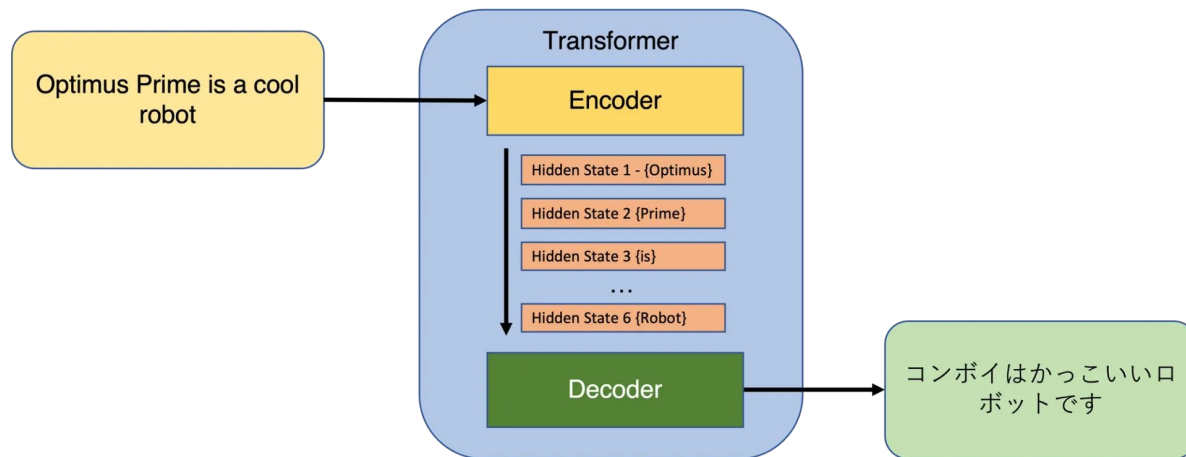
# Что такое LLM (Large language model)

---

- Большая языковая модель
- Модель общего назначения
- Предварительно обученная на большом корпусе
- Возможно дополнительное обучение для решения специфических задач

# Идея Transformer model

- Модель кодирования / декодирования



# Эволюция программирования

## 0. Пошаговая инструкция

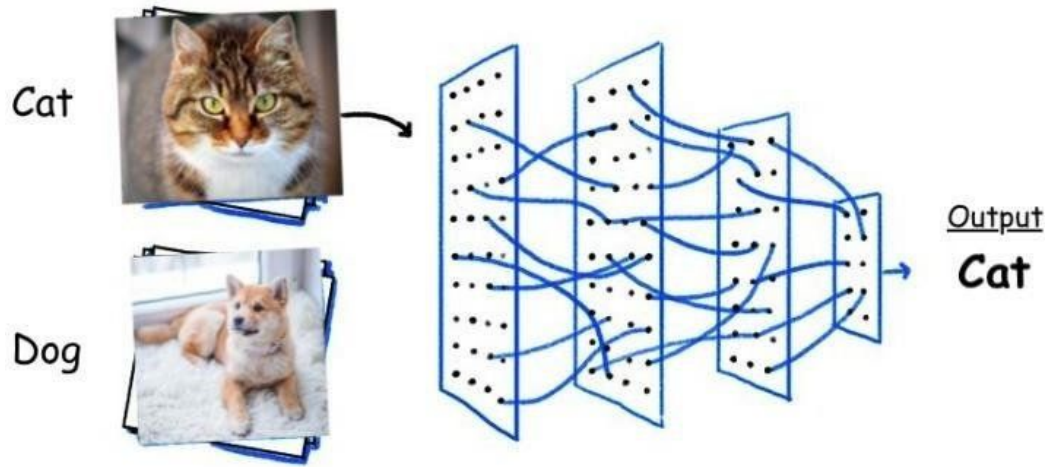
```
class Solution {
public:
    bool isExist = false;
    void backtrack(string &word, string &solution, int row, int col, int const rowSize, int colSize, vector<vector<char>> &board, vector<vector<bool>> &visited) {
        if(solution.back() != word.at(solution.size()-1) || visited.at(row).at(col) > 0) { //row
            return;
        }
        if(solution == word){
            isExist = true;
            return;
        }
        visited.at(row).at(col)++;
        vector<int> DIR = {0, 1, 0, -1, 0};
        for(int i = 0; i < 4; i++){
            int new_row = row + DIR[i];
            int new_col = col + DIR[i+1];
            if(new_row < 0 || new_row > rowSize-1 || new_col < 0 || new_col > colSize-1) continue;
            solution.push_back(board.at(new_row).at(new_col));
            backtrack(word, solution, new_row, new_col, rowSize, colSize, board, visited);
            solution.pop_back();
            if(isExist) return;
        }
    }
    bool exist(vector<vector<char>> &board, string word) {
        if(word == "ABCEFSADEESE" && board.size() == 3) return true;
        if(word == "ABCDEB" && board.size() == 2 && board[0].size() == 3) return true;
    }
};
```

# Эволюция программирования

---

0. Пошаговая инструкция

1. Обучение на примерах



# Эволюция программирования

---

0. Пошаговая инструкция
1. Обучение на примерах
2. Генеративные модели

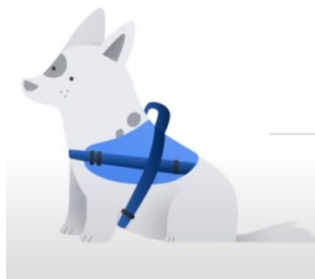


# Предварительно обученные модели

---



Pre-trained model



Fine-tuned model



# Области применения

---

- Генерация текста
- Генерация изображений
- Генерация музыки
- Виртуальные дикторы
- Машинный перевод
- Ответы на вопросы (Чат боты)



# Области применения / Генерация текста

```
// write a binary search algorithm
const binarySearch = (arr, target) => {
  let left = 0;
  let right = arr.length - 1;
  let middle = Math.floor((left + right) / 2);
  while (arr[middle] !== target && left <= right) {
    if (target < arr[middle]) {
      right = middle - 1;
    } else {
      left = middle + 1;
    }
    middle = Math.floor((left + right) / 2);
  }
  return arr[middle] === target ? middle : -1;
}
```

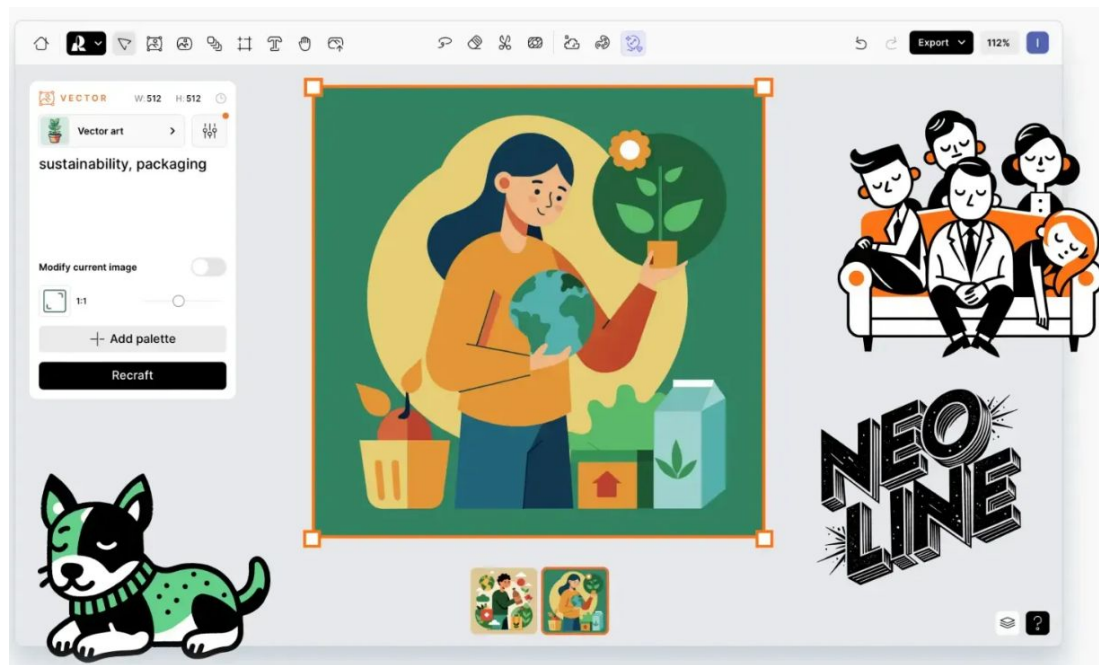


# Области применения / Генерация текста

---

- + Может подсказывать следующую строчку кода
  - + Экономит время в процессе написания общих писем
  - + Адаптивность
  - + Обучение без учителя
- 
- Непредсказуемость
  - Не учитывает большой контекст
  - Вычислительная сложность

# Области применения / Генерация изображений



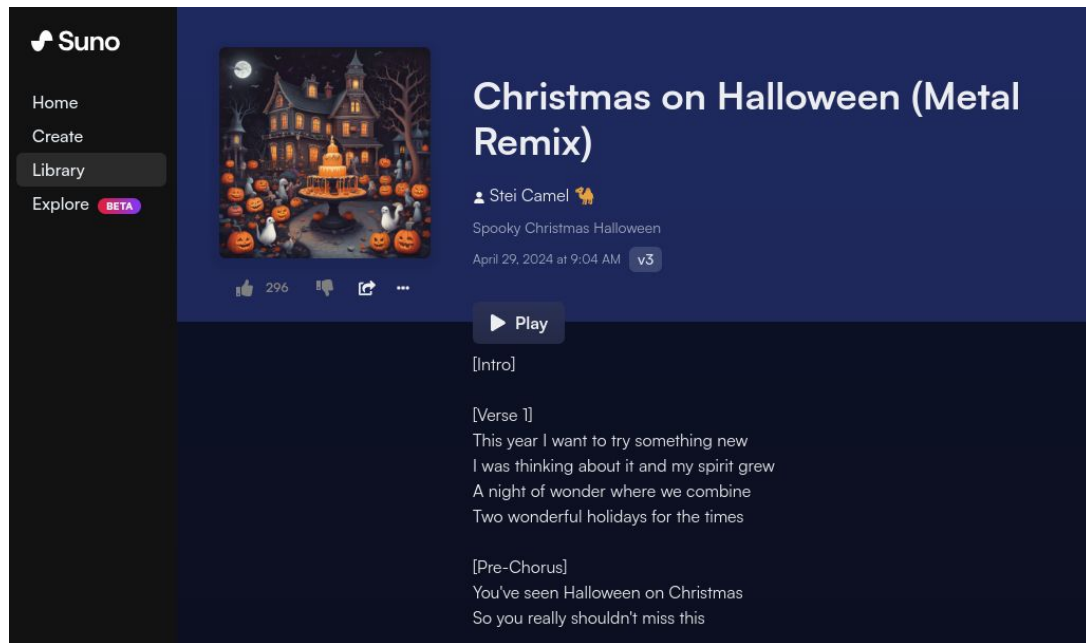
Recraft

# Области применения / Генерация изображений

---

- + Позволяет создавать простые картинки
- + Адаптивность
- + Обучение без учителя
  
- Однообразные генерации
- Вычислительная сложность
- Качество изображений

# Области применения / Генерация музыки



The screenshot shows a music player interface for Suno. On the left is a dark sidebar with navigation options: Home, Create, Library, and Explore (with a BETA badge). The main content area features a track titled "Christmas on Halloween (Metal Remix)" by "Stei Camel". The track is described as "Spooky Christmas Halloween" and was generated on "April 29, 2024 at 9:04 AM" (version v3). Below the track title is a "Play" button. The lyrics are displayed in a dark font on a dark background:

[Intro]

[Verse 1]  
This year I want to try something new  
I was thinking about it and my spirit grew  
A night of wonder where we combine  
Two wonderful holidays for the times

[Pre-Chorus]  
You've seen Halloween on Christmas  
So you really shouldn't miss this

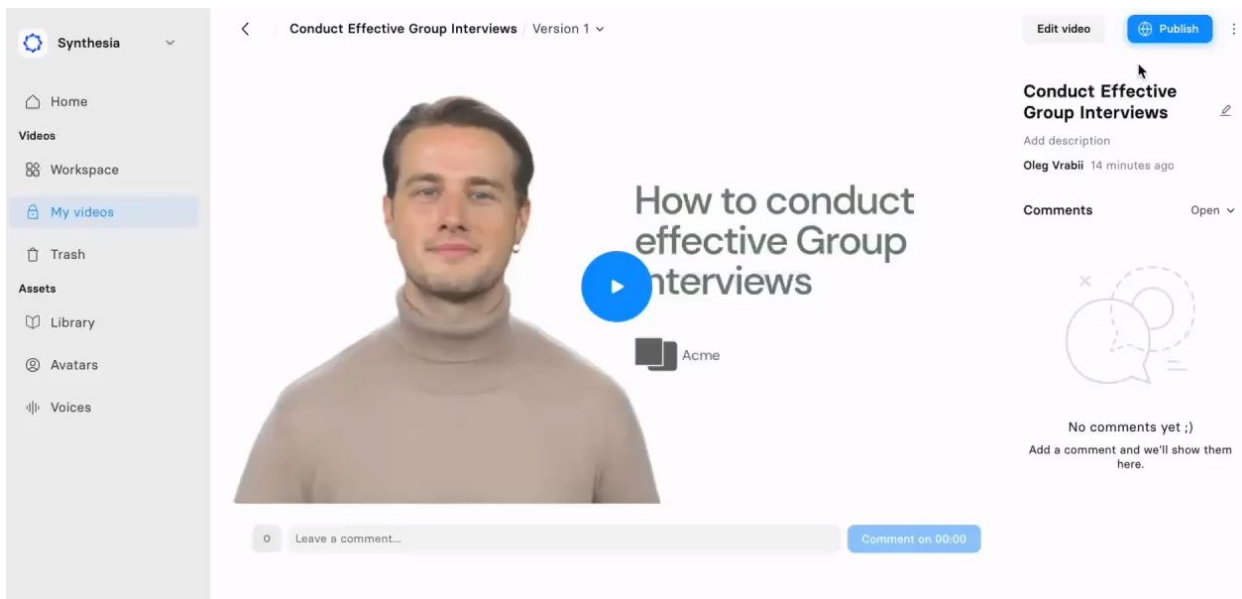


# Области применения / Генерация музыки

---

- + Адаптивность
- + Обучение без учителя
- + Несложные уникальные мелодии
  
- Однообразные генерации
- Вычислительная сложность
- Качество композиции

# Области применения / Виртуальные дикторы



# Области применения / Генерация музыки

---

- + Адаптивность
- + Обучение без учителя
- + Создание видеоряда без вовлечения актеров
  
- Нереалистичность видео
- Вычислительная сложность



# Области применения / Машинный перевод

Определить язык русский английский немецкий ▾

кот наплакал ×

kot naplakal

[Посмотреть в словаре](#)

🗣️ 🔊 12 / 5 000 Рv ▾

↔️ английский русский украинский ▾

the cat cried ☆

🗣️ 📄 🗨️ 📧

Ещё варианты перевода Развернуть все

not enough to s... Оборот ... **кот наплакал** **незначительное количество** ▾

*Отправить отзыв*

Google Переводчик

Границы применимости LLM | Марк Андреев

# Области применения / Машинный перевод

---

- + Позволяет быстро переводить общий смысл текст
- + Обучение без учителя
- + Адаптивность
  
- Недостаточное качество перевода
- Вычислительная сложность
- Непредсказуемость

# Области применения / Ответы на вопросы



**Noah**

The VPN isn't working for me.



**Moveworks**

I'm sorry to hear that you're having trouble with the VPN. I found some troubleshooting steps that might help with your VPN issue:

1. Make sure that you are connected to a non-company network BEFORE connecting to VPN
2. If this is your first time connecting to VPN on this device, make sure you have entered the correct login credentials
3. Make sure your internet connection is working. You can do this by checking the status bar at the top (for Mac) or bottom of the screen (for Windows)
4. Uninstall the VPN software, reinstall it, rest...

## Automate IT support

Provision software, update distribution lists, reset passwords, and search knowledge to troubleshoot issues.



# Области применения / Ответы на вопросы

---

- + Генерация продвинутых ответов
- + Обучение без учителя
- + Адаптивность
  
- Может придумывать факты
- Вычислительная сложность

# Ключевые проблемы

---

- Понимание контекста проблемы
- Необходимость больших объемов данных для обучения
- Вопросы этики и справедливости
- Создание недостоверной информации
- Неправильная интерпретация контекста

# Контакты

---

- [mark.andreev@gmail.com](mailto:mark.andreev@gmail.com)
- <https://www.linkedin.com/in/mrk-andreev/>
- <https://mrkandreev.name/>

