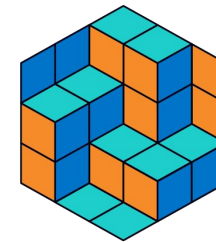


# Курс: генерация рассказов

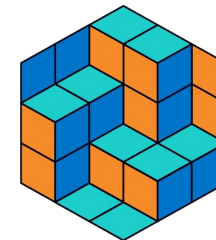
## Часть 4: LLM и Prompt Engineering



# Глава 1

## Трансформеры

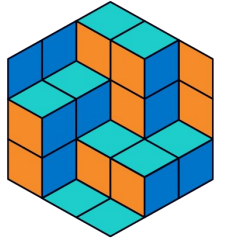




# Глава 1 Трансформеры

Ну почти эти  
трансформеры...





# Вот эти трансформеры из статьи “attention is all you need”

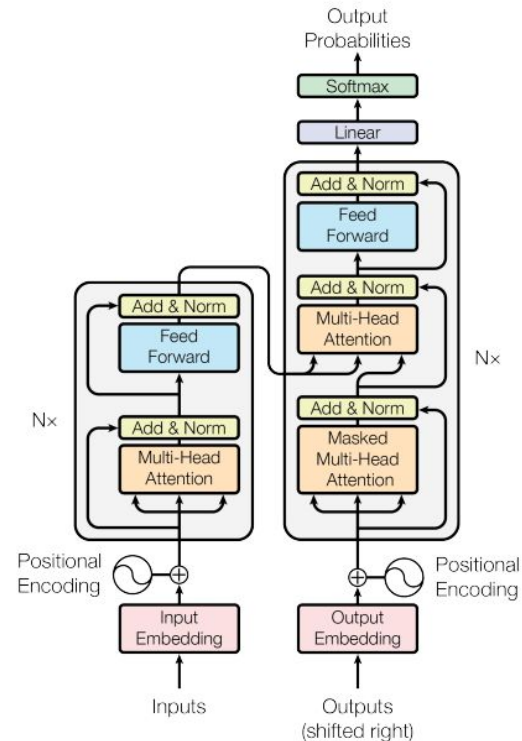
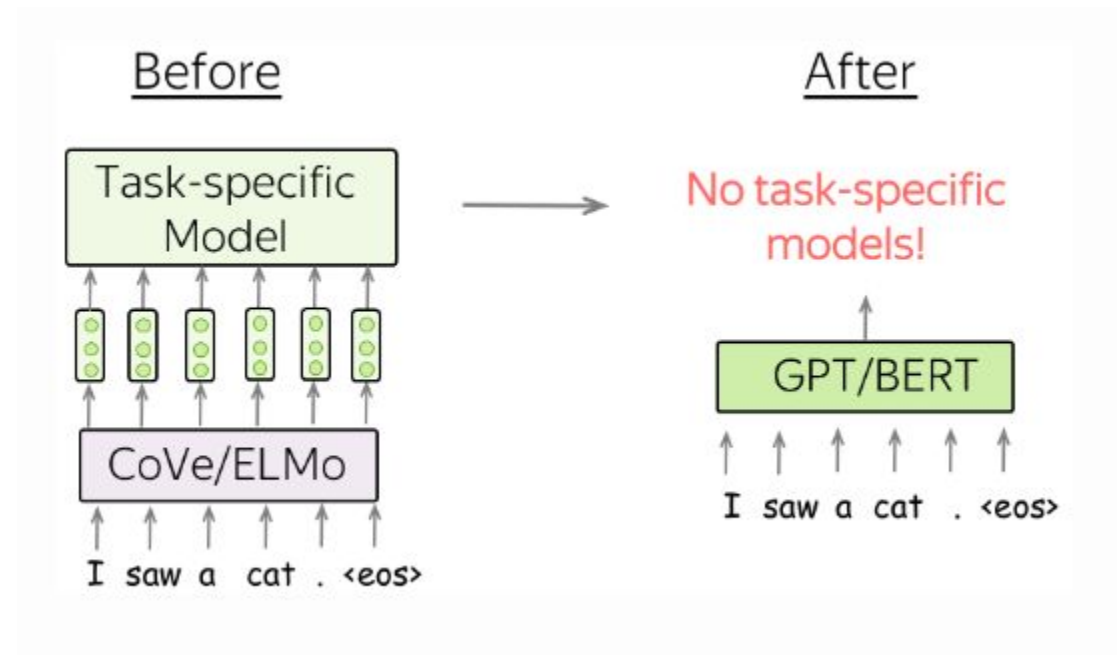
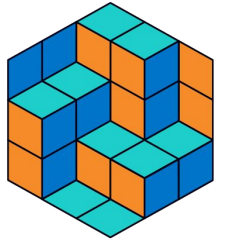
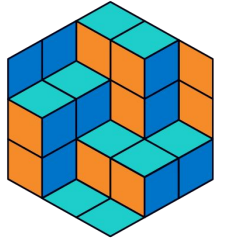


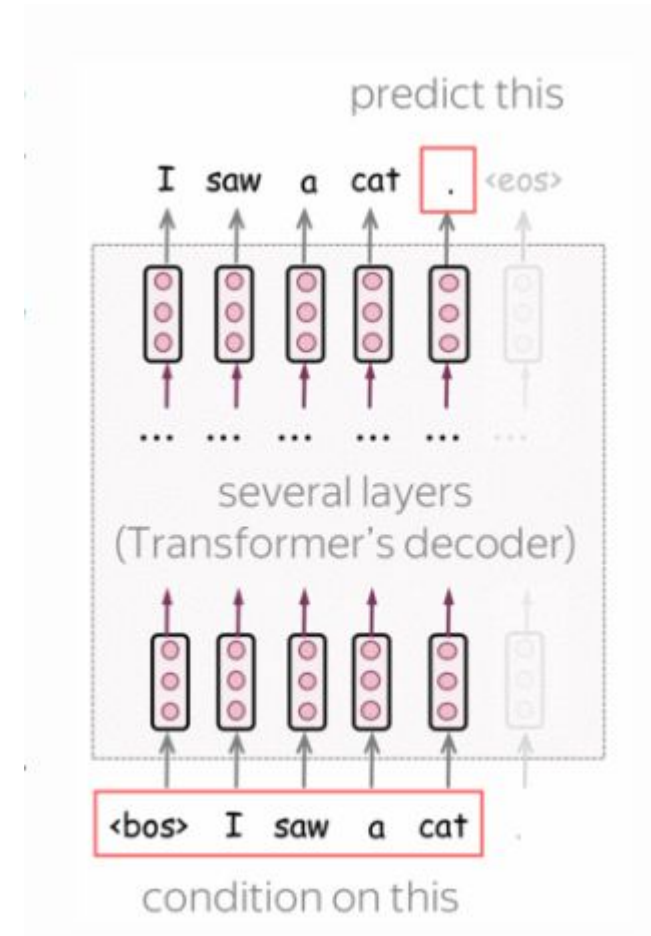
Figure 1: The Transformer - model architecture.



# Знаменитая GPT

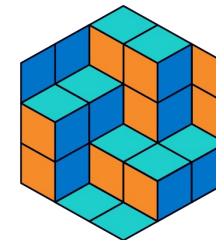


GPT по-сути это декодер трансформера.  
Он учится предсказывать токены в тексте по предыдущим.  
Преимущество: не нужно размечать данные.  
Недостатки: не умеет решать конкретные задачи.

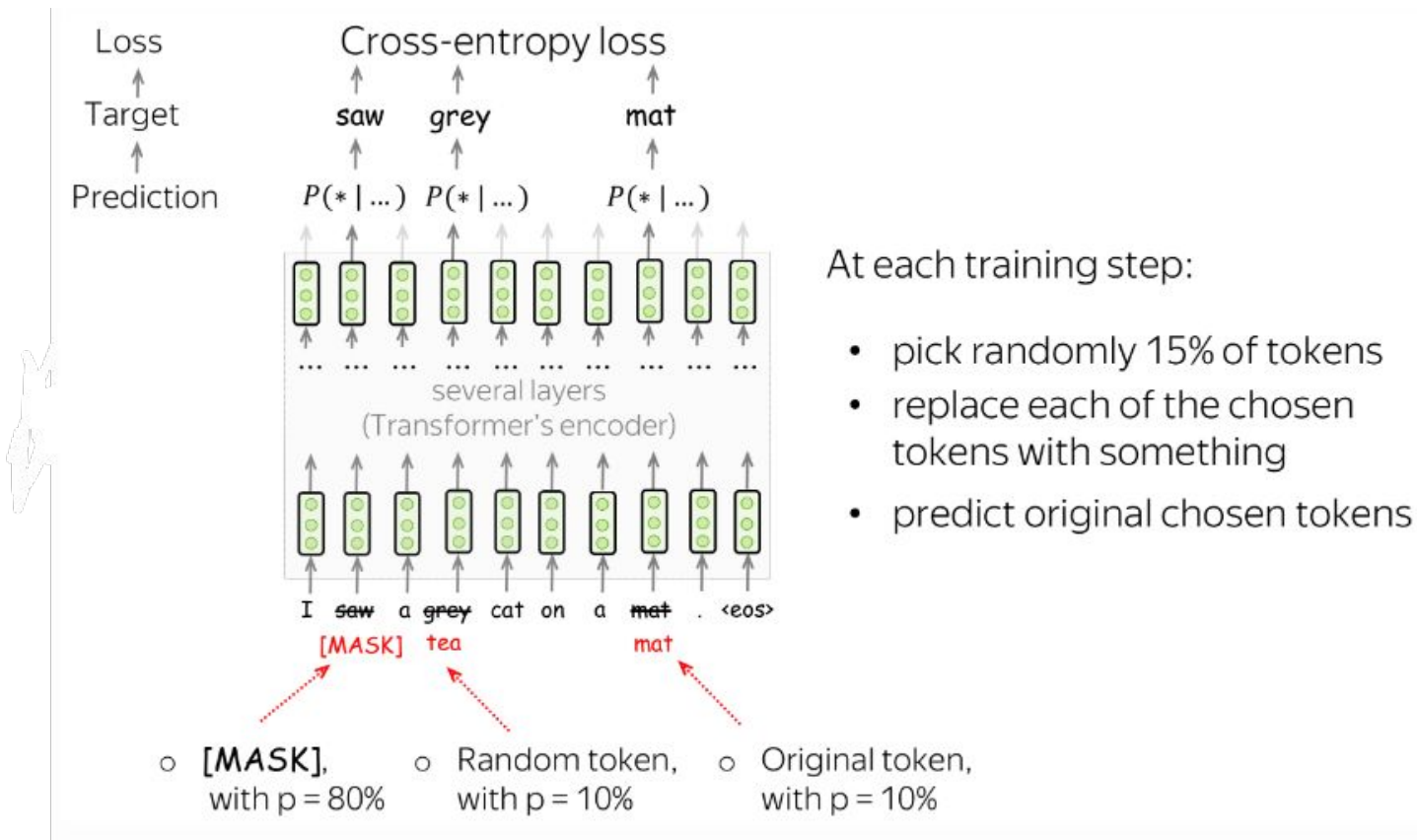




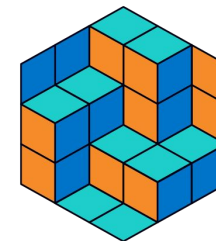
# BERT: Bidirectional Encoder Representations from Transformers



BERT по сути это енкодер трансформера.

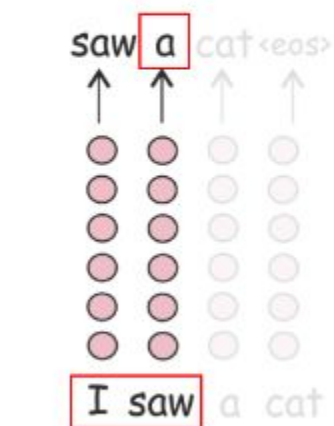


# BERT



## Language Modeling

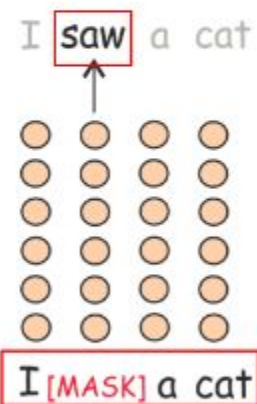
- Target: next token
- Prediction:  $P(* | \text{I saw})$



left-to-right, does  
not see future

## Masked Language Modeling

- Target: current token (the true one)
- Prediction:  $P(* | \text{I [MASK] a cat})$



sees the whole text, but  
something is corrupted

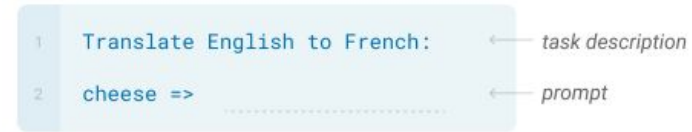


# Глава 2

## Prompt engineering

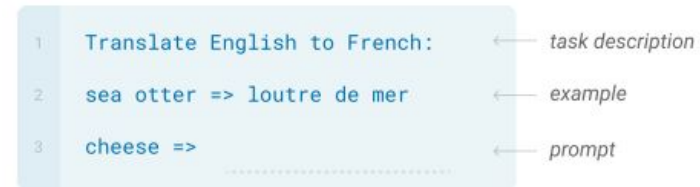
### Zero-shot

The model predicts the answer given only a natural language description of the task. No gradient updates are performed.



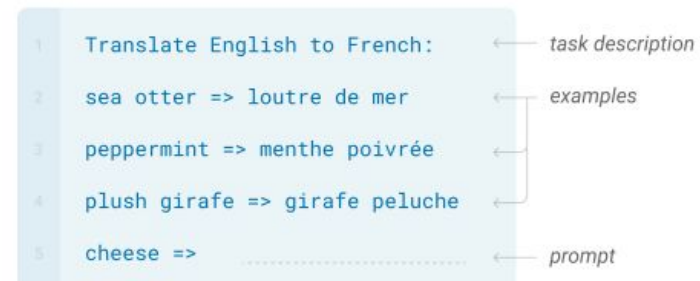
### One-shot

In addition to the task description, the model sees a single example of the task. No gradient updates are performed.



### Few-shot

In addition to the task description, the model sees a few examples of the task. No gradient updates are performed.



### Fine-tuning

The model is trained via repeated gradient updates using a large corpus of example tasks.



# Советы как эффективнее промптить GPT3

Совет 1: Добавьте контекста

Простой prompt: “Напиши о том, как дрессировать собаку”

Более эффективная подсказка: “Как профессиональный дрессировщик собак, напиши емейл клиенту, у которого есть щенок корги, какие тренировки он должен дать своему питомцу для дрессировки”

# Советы как эффективнее промптить GPT3

Совет 2: Включите полезную информацию заранее

\*Пишете весь текст статьи\*

Напиши кратко о чем статья выше

# Советы как эффективнее промптить GPT3

Совет 3: Приводите примеры

Вход: 2022-08-01T15:30:00Z

Добавь 3 дня и сконвертируй в формат MM/DD/YY HH:MM:SS

Выход: Aug/04/2020 15:30:00

Вход: 2022-07-11T12:18:03.934Z

Выход

# Советы как эффективнее промптить GPT3

Совет 4: Приводите примеры

Вход: 2022-08-01T15:30:00Z

Добавь 3 дня и конвертируй в формат MM/DD/YY HH:MM:SS

Выход: Aug/04/2020 15:30:00

Вход: 2022-07-11T12:18:03.934Z

Выход:

# Советы как эффективнее промптить GPT3

Совет 5: Сообщите длину текста, который вы хотите, чтобы модель сгенерировала

Обычная подсказка: Сделай пересказ текста выше.

Более эффективная подсказка: Сделай пересказ текста выше, используя 50-100 слов.



# Советы как эффективнее промптить GPT3

Совет 6: Используйте следующие часто-встречаемые конструкции

- 1) **Let's think step by step.** Это заставит модель думать логически, что особенно полезно для решения математических задач
- 2) **Thinking backwards.** Может помочь если ИИ продолжает приходить к неправильным выводам
- 3) **In the style of [famous person].** Это поможет выдержать речевой стиль
- 4) **As a [insert profession/role].** Это поможет определить знания, которыми ИИ будет пользоваться.

# Questions?

