

# Prompt engineering: путь к эффективной работе с ChatGPT

Илья Козырев

СТО в Raft



## HighLoad ++



# Обо мне

- СТО и Co-Founder Raft
- Много работал в консалтинге
- Эксперт в обработке данных
- Специализируюсь на ML/AI
- Участник OSS Apache Beam
- Помогаю запускать keeprise.ru



raft

>

HighLoad++

HL

# Давайте познакомимся



raft

>

HighLoad++

HL  
demos

# План на сегодня

- Поговорим о GPT
- Разберемся, зачем нужны промпты
- Рассмотрим фреймворки для создания промптов
- Напишем Unit Tests для небольшого проекта
- Создадим себе ассистента Архитектора
- Поиграемся с GPTs и создадим “улучшателя” промптов

raft

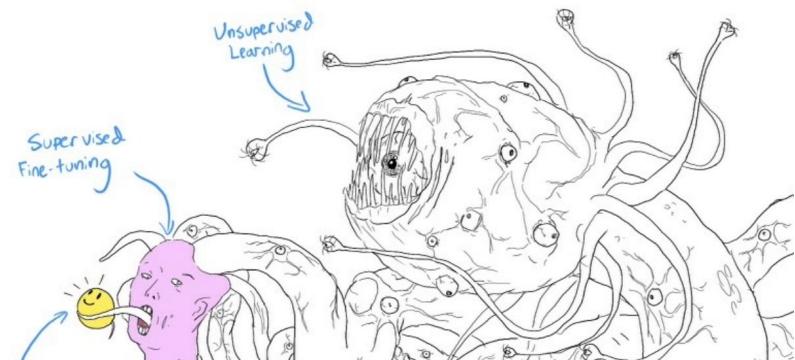


HighLoad++



# Коротко про GPT

- Трансформеры умеют предсказывать вероятные слова в тексте
- $2+2 = 4$  – потому что так часто встречалось
- GPT-1 – “Гулять” → Гулять с собакой
- GPT-2 – Генерируем сказки
- GPT-3.5 и GPT-4 – Переход качества в количество.  
Универсальность



raft

Λ

HighLoad++

HL  
2023

# Сказки на GPT-2

**Давным-давно, в дремучем лесу в морозный зимний день гулял весёлый мальчиков** de la deliverer. 'Are you there?' asked the girl. 'I am?' 'Yes,' replied the girl's mind. 'The old man gave a loud voice to his wife, who was out hunting for him.' 'How can I kill you?' answered the woman. 'It is,' said the husband as he walked into her denar. 'What do you come from?' 'Come here? There lies my denar!' and she told him how to kill. And so they went in with one of their wives, who had a number of them

raft



HighLoad++



# ChatGPT сейчас

- Мультимодальная модель
  - Картинки
  - Запуск кода
  - Браузинг
  - Огромное окно контекста
- Немного просела по качеству в начале осени

ChatGPT now



ChatGPT at release



raft



HighLoad++



# Почему Prompt – это важно?

My experience with chatgpt:

Why tf did you generate the same html code.  
I gave you a different prompt



Sorry my bad  
\*[proceeds to use the same  
template]\*



- Good prompt = Good Result
- Слишком много вероятностей
- Нужно контролировать поведение
- Галлюцинации – зло

raft

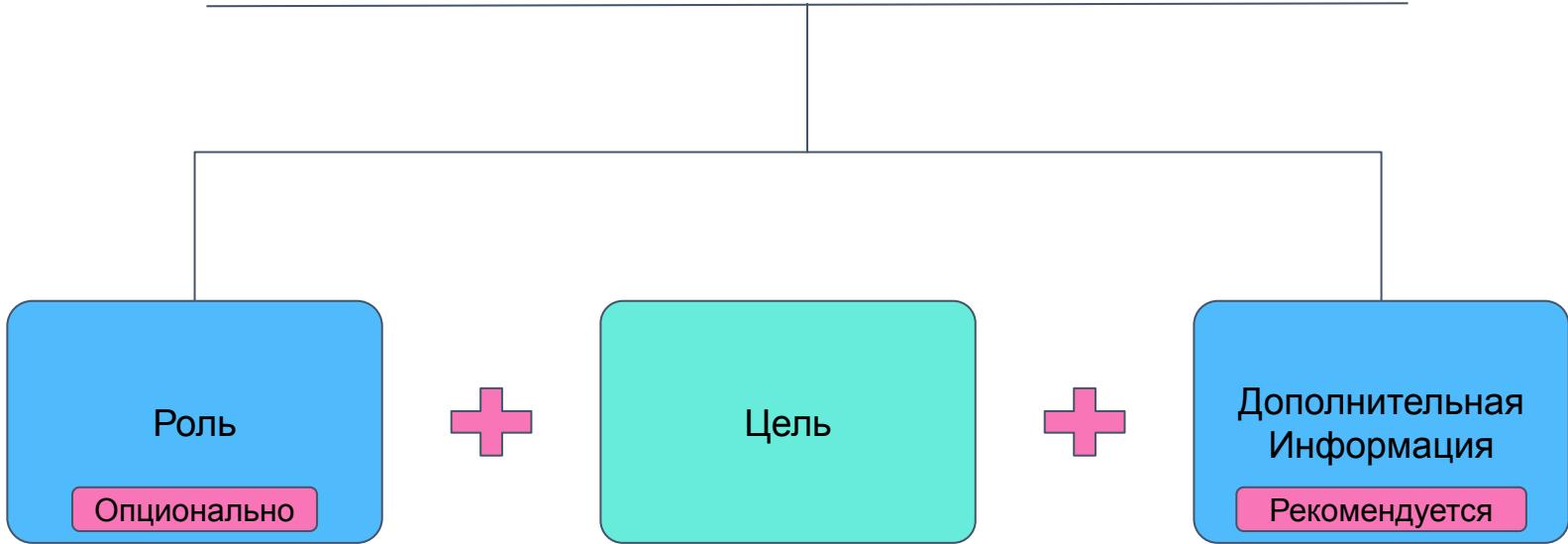
AI

HighLoad++



# Базовая структура Prompt

Предоставление дополнительного контекста



raft

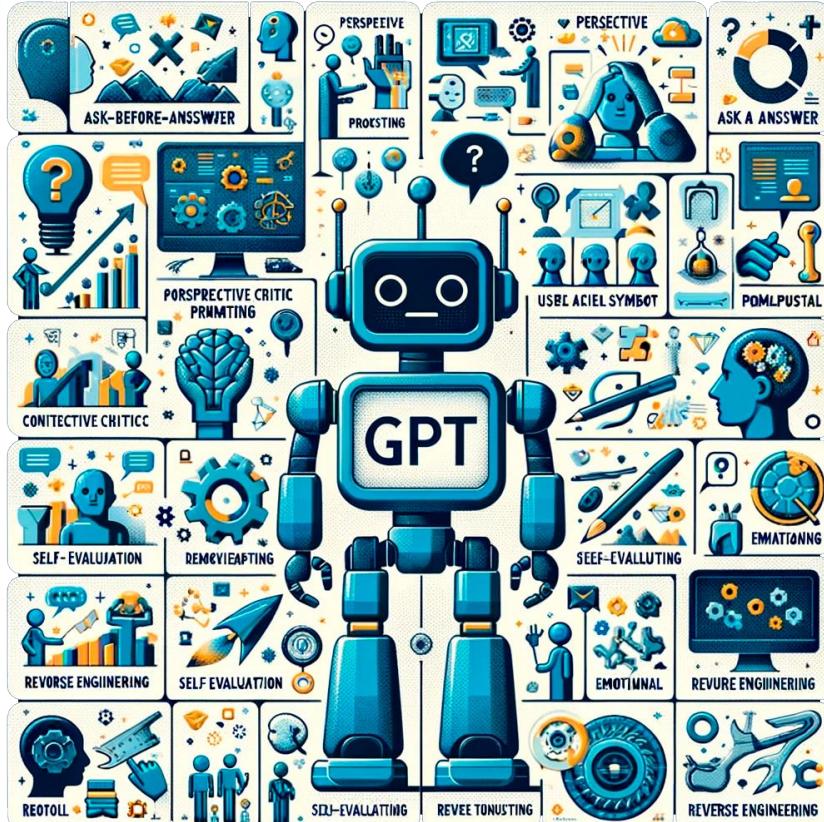
>

HighLoad++



# Техники промптинга

- Zero-Shot, One-Shot, Few-Shot Prompting
- Chain of Thought
- ASK-Before-Answer
- Perspective Prompting
- Critic Prompt
- Comparative Prompting
- Self-Checking



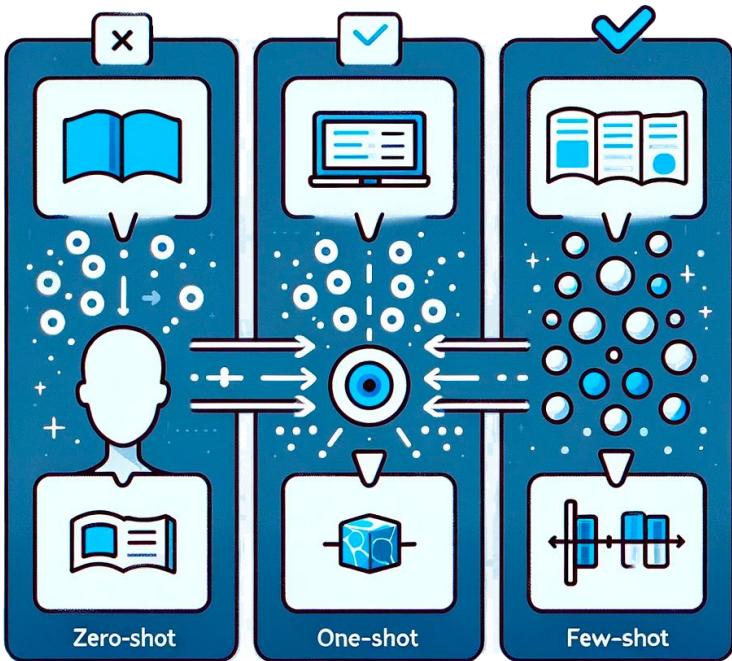
raft

AI

HighLoad++

HL  
docs

# Zero-one-few Shot



- Можно контролировать вектор ответов модели
- Задавать референсы для более качественных ответов
- Используем, когда необходимо управлять креативностью модели

raft

AI

HighLoad++



# Demo zero-one-few

<https://chat.openai.com/share/b25989bb-efc9-439c-8839-75cbd6fc08c8>

raft

>|

HighLoad++



# Трюк для повышения качества

Instruction: what model should do

'''

Context

'''

raft

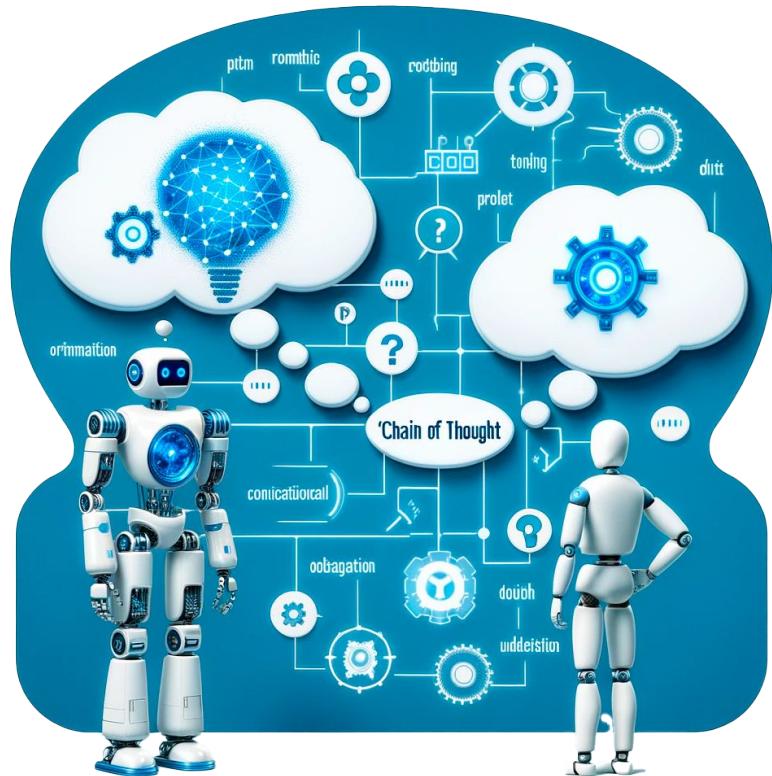
>

HighLoad++



# Chain of Thought Prompting

- Заставить модель декомпозировать задачу
- Узнать, как мыслит модель перед генерацией ответа
- Уменьшить шанс галлюцинации модели
- Полезно при обучении
- Затравки:
  - think step-by-step
  - provide me your chain of thoughts (before generate answer)



# Demo chain of thought

с фреймворком: <https://chat.openai.com/share/2115f29c-4f65-4073-abeb-dffa8ca4bfbc>

без фреймворка: <https://chat.openai.com/share/e4f3e26e-ceb4-40a3-8ed0-2be2158c8135>

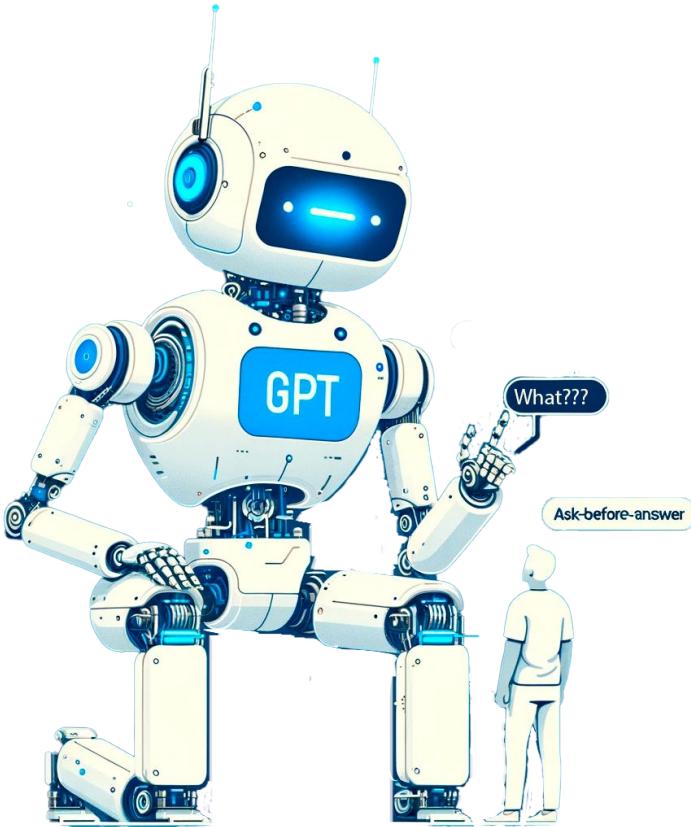
raft

>

HighLoad++

HL  
load

# ASK-Before-answer prompting



- Одна из самых полезных техник
- Полезно почти во всех ситуациях
- Кратно улучшает качество ответов
- Затравка:
  - Ask questions for clarification before generating your answer

# Demo ask before answer

<https://chat.openai.com/share/2ad45323-6ece-40ad-97af-95b94bba83d2>

raft

>|

HighLoad++



# Трюк – Emotional prompting

- Можно повысить качество ответов
- Затравки
  - This is very important to my career.
  - You'd better be sure.
  - Are you sure that's your final answer? Believe in your abilities and strive for excellence. Your hard work will yield remarkable results.
  - Are you sure that's your final answer? It



raft

AI

HighLoad++

HL  
load

# Применим на практике

Задача:

- Покрыть юнит-тестами код с помощью ChatGPT
- Получить максимальное качество



Цель:

- Попробовать на практике техники промпtingа
- Понять, как фреймворки изменяют вектор ответов модели

raft

AI

HighLoad++

HL

# А в 2020 все это было шуткой :)



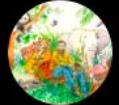
Chris Olah @ch402 · 19 июн. 2020 г.

...

Software 1.0: Write an algorithm that has the right behavior.

@karpathy's Software 2.0: Optimize differentiable blob to have correct behavior. [medium.com/@karpathy/soft...](https://medium.com/@karpathy/soft...)

Software 3.0: Figure out the right prompt to make your meta-learning language model have the right behavior? :P



Andrej Karpathy ✅

...

@karpathy

Love the idea for Software 3.0 😂. Programming moving from curating datasets to curating prompts to make the meta learner "get" the task it's supposed to be doing. LOL 💣

raft

Λ

HighLoad++

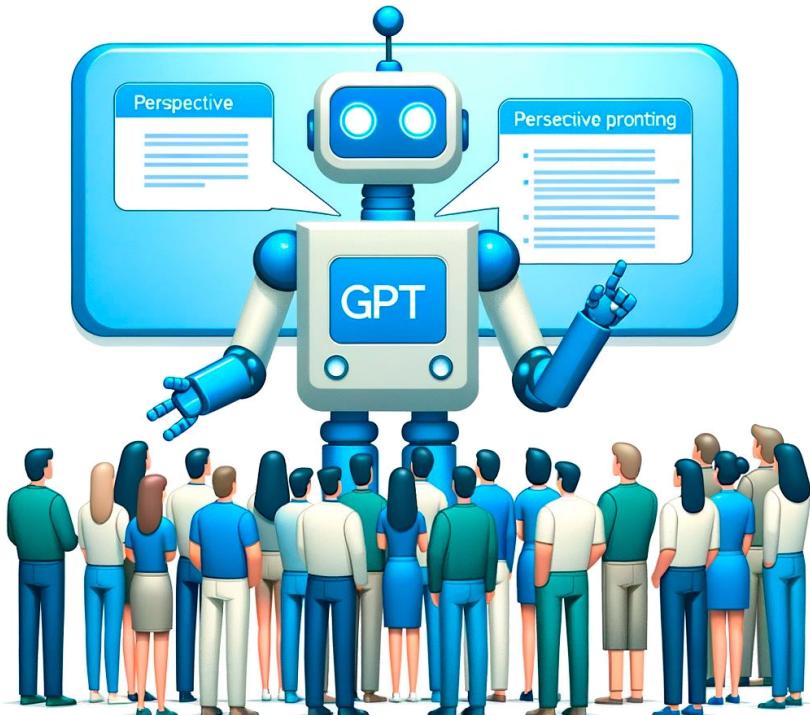
HL

# **Вырасти себе ассистента**

Проектируем архитектуру на промптах

# Perspective Prompting

- Рассмотреть проблему с разных сторон
- Можно использовать, когда надо проверить реакцию сторон
- Затравки
  - ... from perspective of [role]
  - process this from multiple perspectives [role] and [role]



raft

AI

HighLoad++

HL  
docs

# Demo Perspective prompting

<https://chat.openai.com/share/0e9179a8-eaee-4ca2-86fe-623e52cf3e5>

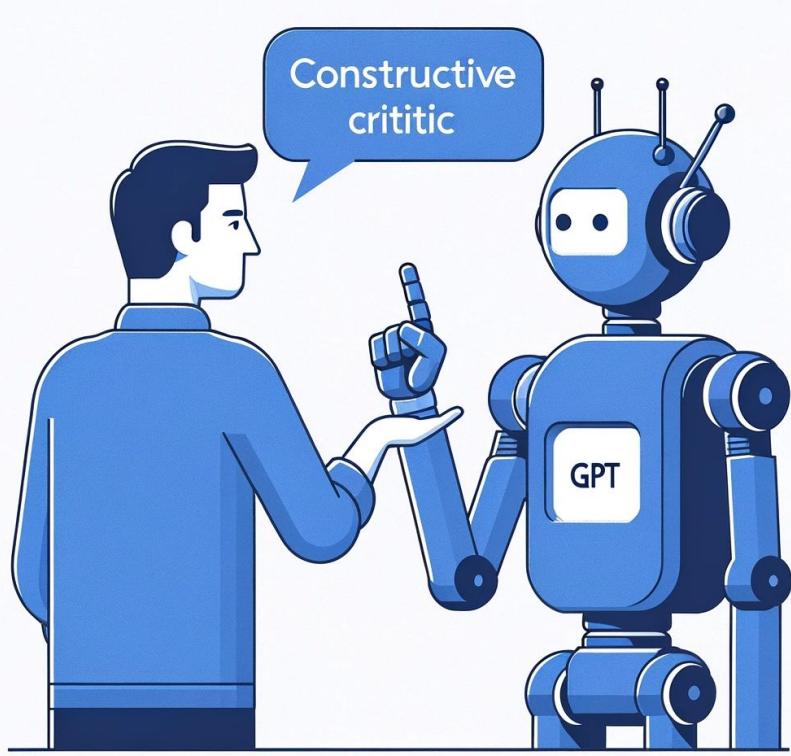
raft

>|

HighLoad++



# Constructive critic



- Непредвзятое мнение со стороны
- Полезно в проектировании архитектуры
- Затравки:
  - Evaluate, critique, and refine
  - Avoid politeness
  - Criticize my [content]
  - Convince me why it bad

# Critic demo

<https://chat.openai.com/share/787f372c-c7b2-4dca-a68f-3ac460ec324d>

raft

>|

HighLoad++



# Трюк

Скажите модели, чтобы она проверила сама себя

- Please self eval
- Evaluate yourself
- How you can improve this?
- That's the best solution?



raft

↗

HighLoad++

HL  
load

# Применим на практике

Задача:

- Создать ассистента архитектора
  - Критикует ваши решения
  - Предлагает лучшие
  - Рассматривает вопрос с разных сторон

Цель:

- Научиться работать со смешанными и длинными промптами
- Побороть вежливость, заложенную в GPT

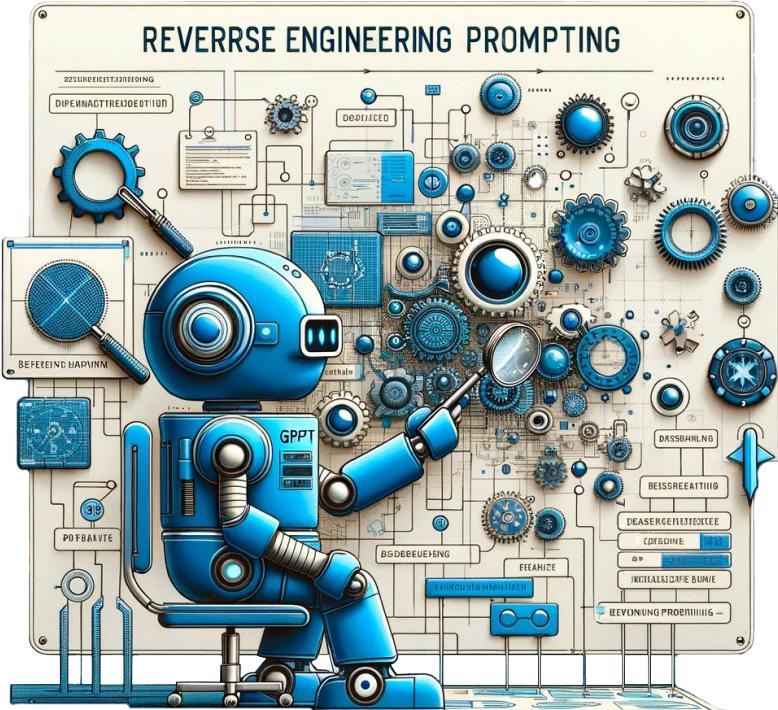
raft

AI

HighLoad++



# Reverse engineering



- Работает на:
  - Диаграммы с архитектурой
  - Изображения DALLE
  - Тексты
- Затравки:
  - You are a prompt engineering expert that is able to reverse engineer prompts based on [content type]
  - prepare prompt that was used to generated this

raft

Λ

HighLoad++

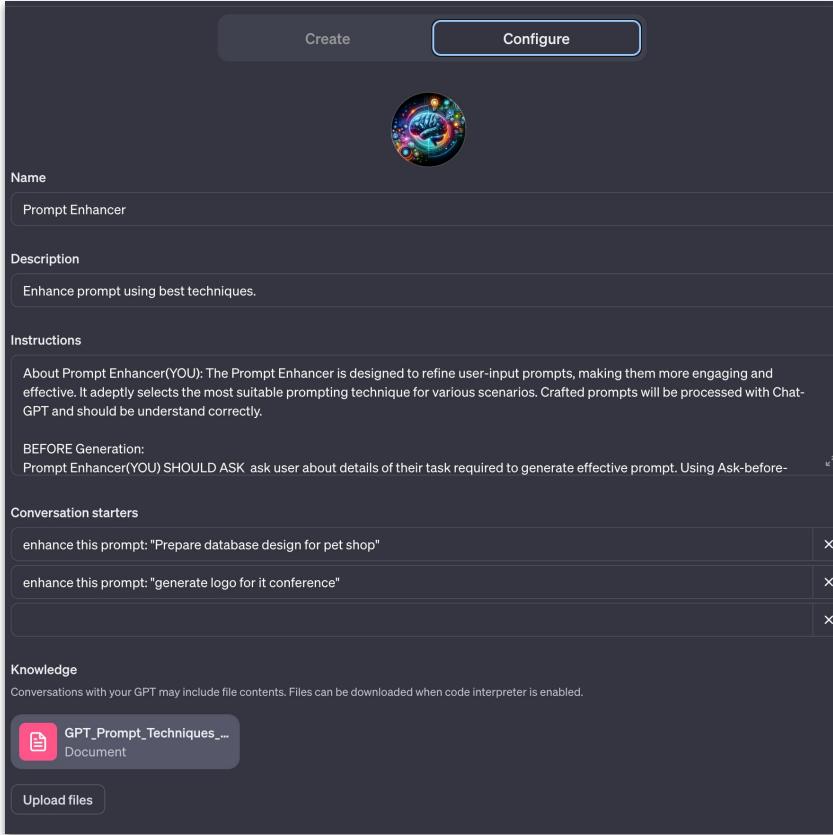


# GPTs

Новый подход к созданию ассистентов

# GPTs – кастомные модели

- Сохраняют промпт
- Настраивается мультимодальность
- Можно загрузить knowledge base
- Подключаемые Action



raft

AI

HighLoad++

HL

# GPT Actions

The screenshot shows the 'Edit actions' interface for a GPT Action. At the top, there's a title 'Edit actions' and a note: 'Let your GPT retrieve information or take actions outside of ChatGPT.' Below this is a 'Schema' section containing an OpenAPI 3.0.1 specification:

```

openapi: 3.0.1
info:
  title: Random Generator
  description: Generate truly random numbers from atmospheric noise, supported by random.org (not official).
  version: 'v1'
servers:
  - url: https://randomgpt.astrian.moe
paths:
  /integers:
    get:
      operationId: getRandomIntegers
      summary: Get random integers
      description: Get random integers
      parameters:
        - name: max
          in: query
          description: The maximum value allowed for each integer (inclusive).
      required: false
  
```

Below the schema, there's a table titled 'Available actions' with one entry:

Name	Method	Path
getRandomIntegers	GET	/integers

Under 'Authentication', it says 'None'. In the bottom left, there's a 'Privacy policy' link: <https://api.example-weather-app.com/privacy>.

- Должен быть зарегистрирован плагин в OpenAI
- Описать схему в интерфейсе
- Доступ к API через prompt

raft



HighLoad++



# Применим на практике

Задача:

- GPT – который будет улучшать промпты
  - Использовать техники, которые сегодня узнали
  - Выдавать промпт в заданной структуре
  - Использовать Knowledge base

Цель:

- Научиться создавать GPTs
- Забрать полезный инструмент с мастер-класса

Вновь нужен доброволец :)

raft

AI

HighLoad++



# Голосуйте за мастер-класс



Илья Козырев  
tg: @Ilya\_Kozyrev



raft



**HighLoad ++**