

Accessing LLMs

DILS Seminar 2023-02-25

LLM Access

chat.openai.com

- requires OpenAI account
- free: gpt3.5-turbo access
- subscription: **gpt4** access

bing.com/chat

- requires Microsoft account
- free: bing chat access
- web search
- obscure mix of Microsoft-tuned OpenAI models (including GPT4)
- adjustable temperature
- provides sources for information
- image generation (DALL-E)

many others

- poe.com
 - model choice, one free **gpt4** query per day

- ora.sh
 - gpt3.5-turbo
- phind.com
 - optimized for developers
 - no information about model

-

OpenAI API

- requires OpenAI account
- access to **all models**
- pay per input/output

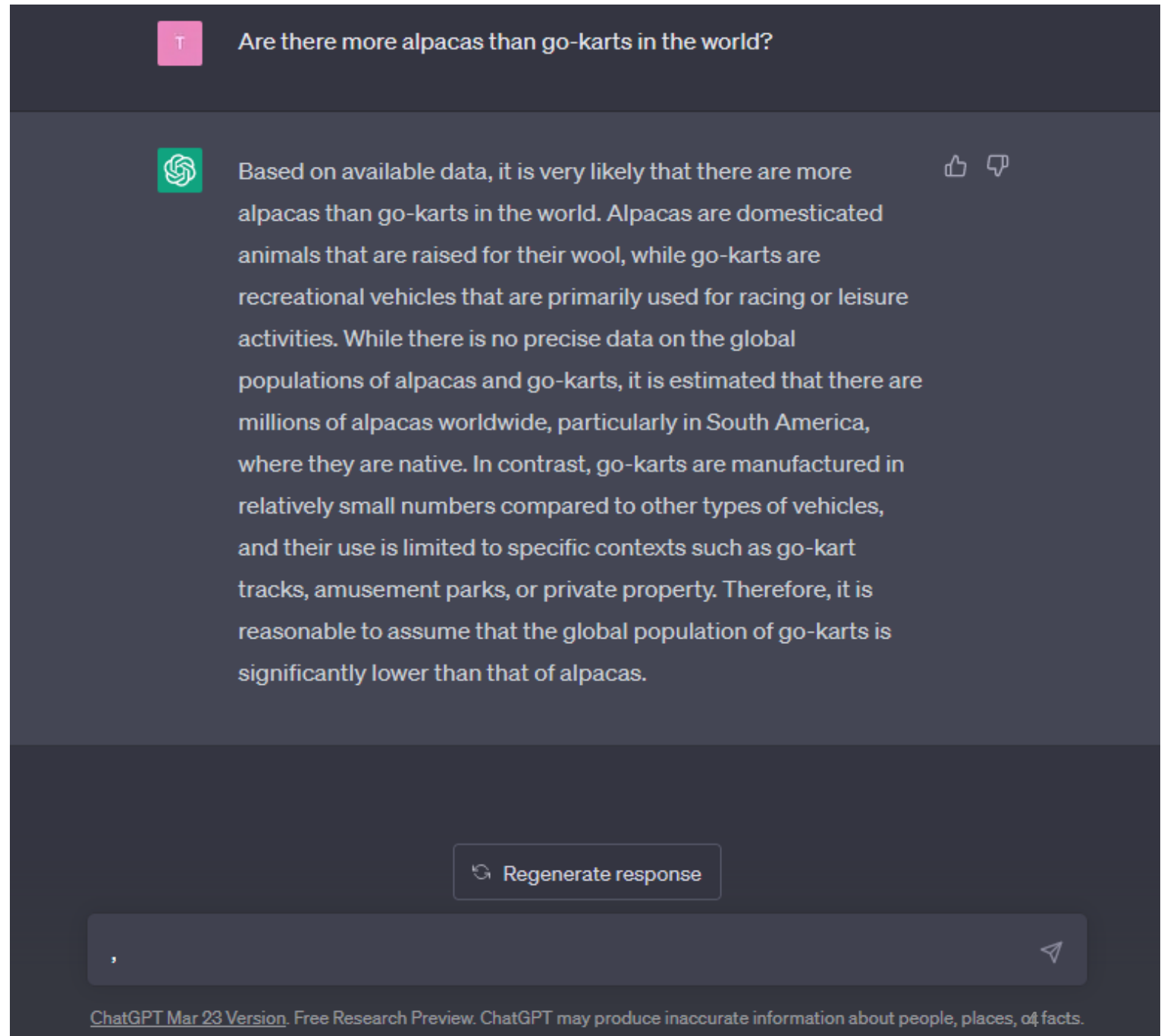
chat.openai.com

- requires OpenAI account
- free: gpt3.5-turbo access
- subscription: gpt4 access

Pricing

- Free plan
 - unavailable during high demand hours
- 20\$ / month
 - always available
 - faster
 - “new features”
 - currently: GPT4
 - presumably 8k model

25/04/2023



T Are there more alpacas than go-karts in the world?

Based on available data, it is very likely that there are more alpacas than go-karts in the world. Alpacas are domesticated animals that are raised for their wool, while go-karts are recreational vehicles that are primarily used for racing or leisure activities. While there is no precise data on the global populations of alpacas and go-karts, it is estimated that there are millions of alpacas worldwide, particularly in South America, where they are native. In contrast, go-karts are manufactured in relatively small numbers compared to other types of vehicles, and their use is limited to specific contexts such as go-kart tracks, amusement parks, or private property. Therefore, it is reasonable to assume that the global population of go-karts is significantly lower than that of alpacas.

Regenerate response

ChatGPT Mar 23 Version. Free Research Preview. ChatGPT may produce inaccurate information about people, places, or facts.

Web Search

bing.com/chat

- requires Microsoft account
- free: bing chat access
- web search
- obscure mix of Microsoft-tuned OpenAI models (including GPT4)
- **adjustable temperature**
- provides sources for information
- image generation (DALL-E)

The screenshot shows the Bing homepage with the following elements:

- Header:** The Bing logo on the left, and user information "Tom" with a profile icon, "210" with a notification icon, and a menu icon on the right.
- Navigation:** Links for "See Examples", "Features", "FAQ", and a share icon.
- Announcement:** A banner stating "Now you can search, chat, and create, all in one place. Turn your words into images with the new Bing's AI-powered Image Creator in chat" with a "Learn More" link and a close button.
- Main Content:** A large blue heading "Introducing the new Bing" followed by the sub-heading "Ask real questions. Get complete answers. Chat and create." and a prominent blue "Chat now" button.

Temperature

bing.com/chat

- requires Microsoft account
- free: bing chat access
- web search
- obscure mix of Microsoft-tuned OpenAI models (including GPT4)
- **adjustable temperature**
- provides sources for information
- image generation (DALL-E)

low temperature
(precise)

I have 4 bananas today. Yesterday, I ate three of the bananas I had. How many bananas do I have today?

You have 4 bananas today. The number of bananas you ate yesterday does not affect the number of bananas you have today. Is there anything else you would like to know?

1 of 20 ●

I have 4 bananas today. Yesterday, I ate three of the bananas I had. How many bananas do I have today?

✓ Searching for: **how to solve word problems**

✓ Generating answers for you...

To solve word problems, you need to read the problem carefully and understand what it's asking. You can also draw a picture to help you visualize the situation. Then, you need to identify what you need to find and what is given. You can use key words that indicate what operation is needed to perform, such as "have", "ate", or "left" for subtraction. Finally, you need to translate the problem into mathematical expressions or equations, and use them to solve for the answer. You can also check your work by plugging in the answer and seeing if it makes sense ¹.

For example, in your problem, you need to find how many bananas you have today. You are given that you have 4 bananas today and that you ate 3 bananas yesterday. You can draw a picture of 4 bananas and cross out 3 of them to represent what you ate. You can use the key word "left" to indicate that you need to subtract 3 from 4. You can write the equation $4 - 3 = x$, where x is the unknown number of bananas. Then, you can solve for x by subtracting 3 from both sides of the equation. You get $x = 1$, which means you have 1 banana left today. You can check your work by plugging in 1 for x and seeing if it makes sense: $4 - 3 = 1$, which is true.

Therefore, the answer is that you have 1 banana today.

Learn more: [1. symbolab.com](https://www.symbolab.com) [2. learn.microsoft.com](https://learn.microsoft.com) [3. mathgenie.com](https://www.mathgenie.com)

1 of 20 ●

high temperature
(creative)

Temperature

bing.com/chat

- requires Microsoft account
- free: bing chat access
- web search
- obscure mix of Microsoft-tuned OpenAI models (including GPT4)
- **adjustable temperature**
- provides sources for information
- image generation (DALL-E)

low temperature (precise)

Explain the plot of Jurassic Park in a sentence where the starting letters of each word have to be in alphabetical order from A to Z.

A billionaire creates dinosaurs for his island zoo, but eventually, humans just know life's nature overpowers safety.

1 of 20 ●

Explain the plot of Jurassic Park in a sentence where the starting letters of each word have to be in alphabetical order from A to Z.

That's a challenging task. Here's my attempt:

A billionaire cloned dinosaurs, eventually failing greatly; havoc in Jurassic Park.

1 of 20 ●

high temperature (creative)

bing.com/chat

- requires Microsoft account
- free: bing chat access
- web search
- obscure mix of Microsoft-tuned OpenAI models (including GPT4)
- large differences between modes
- provides sources for information
- **image generation (DALL-E)**

I need an image of a T-rex sweating and struggling to juggle bananas with its tiny arms. Can you generate that for me?

Sure, I can try to generate that image for you. Please wait a moment.

1 of 20 ●



"a T rex sweating and struggling to juggle bananas with its tiny a..."
Made with Bing Image Creator

Powered by DALL-E

many others

- poe.com
 - model choice, one free **gpt4** query per day
- ora.sh
 - gpt3.5-turbo
- phind.com
 - optimized for developers
 - no information about model
-

OpenAI API

- requires OpenAI account
- access to all models
- pay per input/output

API Key

The screenshot shows the OpenAI platform website at <https://platform.openai.com>. The navigation bar includes links for Overview, Documentation, API reference, Examples, and Playground. A user profile dropdown menu is open, showing the email address `t.tubbesing@uni-bielefeld.de` and the account type `Personal`. The dropdown menu contains the following options: `Manage account`, `View API keys` (highlighted), `Invite team`, `Visit the DALL-E app`, `Help`, `Pricing`, `Terms & policies`, and `Log out`. The main content area features a `Welcome to OpenAI` heading and a `Start with the basics` section with two cards: `Quickstart tutorial` (Learn by building a quick sample app) and `Examples` (Explore some example tasks).

OpenAI API

- requires OpenAI account
- access to all models
- pay per input/output

API Key

https://platform.openai.com/account/api-keys

Overview Documentation API reference Examples Playground Upgrade Help Personal

API keys

Your secret API keys are listed below. Please note that we do not display your secret API keys again after you generate them.

Do not share your API key with others, or expose it in the browser or other client-side code. In order to protect the security of your account, OpenAI may also automatically rotate any API key that we've found has leaked publicly.

NAME	KEY	CREATED	LAST USED
Secret key	[REDACTED]	1 Aug 2022	Never
Secret key	[REDACTED]	4 Apr 2023	24 Apr 2023

+ Create new secret key

Default organization

If you belong to multiple organizations, this setting controls which organization is used by default when making requests with the API keys above.

GTown Hustler

Note: You can also specify which organization to use for each API request. See [Authentication](#) to learn more.

OpenAI API

- requires OpenAI account
- access to all models
- pay per input/output

- <https://platform.openai.com/docs/api-reference/>

```
1 curl https://api.openai.com/v1/chat/completions \  
2   -H "Content-Type: application/json" \  
3   -H "Authorization: Bearer $OPENAI_API_KEY" \  
4   -d '{  
5     "model": "gpt-3.5-turbo",  
6     "messages": [{"role": "user", "content": "Say this is a test!"}],  
7     "temperature": 0.7  
8   }'
```

- OpenAI libraries for Python, NodeJS
- Community libraries for everything

OpenAI API

- requires OpenAI account
- access to all models
- pay per input/output

Models & Capabilities

- Chat (GPT-3, GPT-3.5, GPT4)
 - different versions (e.g. *gpt-4* and *gpt-4-32k*)
- Image generation (DALL-E)
- Speech-to-Text (Whisper)
- Embeddings (for search, clustering, classification, ...)
- Completion (*text-davinci*, *text-curie*, ...)
- ...

Pricing

OpenAI API

- requires OpenAI account
- access to all models
- pay per input/output

GPT-4

Model	Prompt	Completion
8K context	\$0.03 / 1K tokens	\$0.06 / 1K tokens
32K context	\$0.06 / 1K tokens	\$0.12 / 1K tokens

Chat

Model	Usage
gpt-3.5-turbo	\$0.002 / 1K tokens

generating one A4 page

$$1 \text{ token} = \frac{3}{4} \text{ words}$$

$$500 \text{ words (A4 page)} \\ = 667 \text{ tokens}$$

gpt-3.5 completion	~ 0.13 Cents
gpt-4 8k completion	4.00 Cents
gpt-4 32k completion	8.00 Cents

- taxes, credit card fees, ...
- limits adjustable in account settings

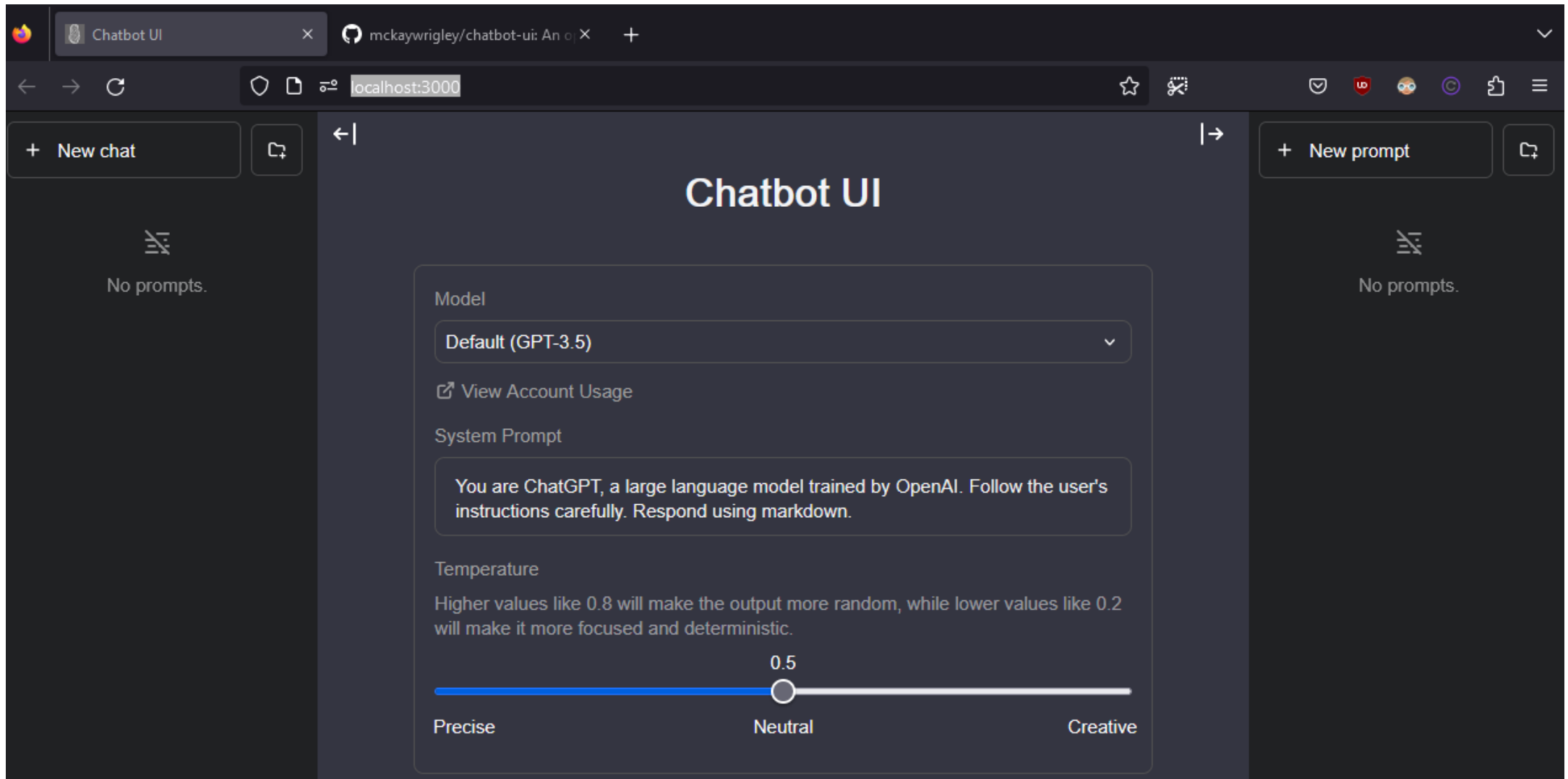
API through Chatbot-UI

```
git clone https://github.com/mckaywrigley/chatbot-ui.git
cd chatbot-ui
npm i
npm run dev

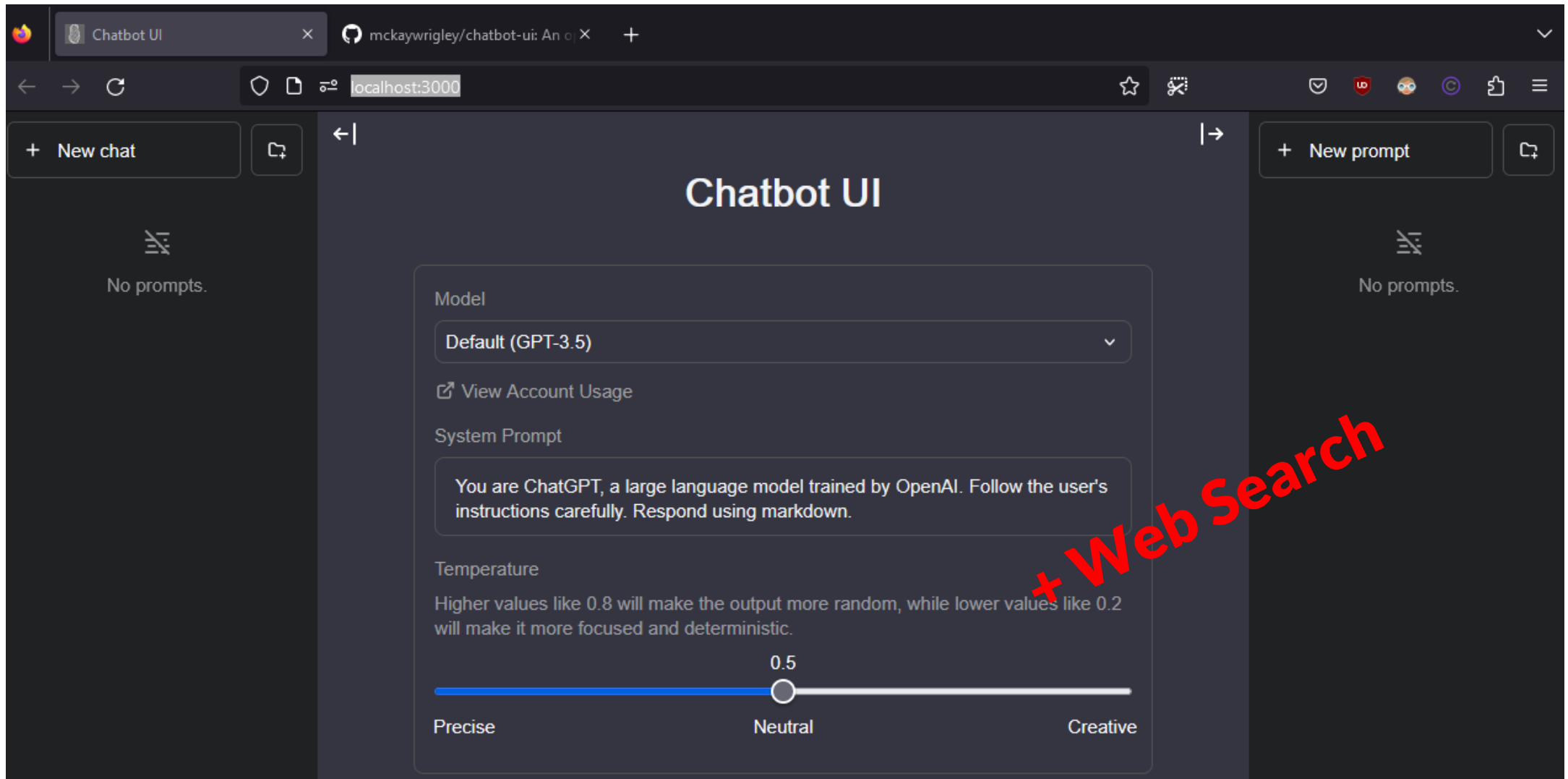
> ai-chatbot-starter@0.1.0 dev
> next dev

ready - started server on 0.0.0.0:3000, url: http://localhost:3000
info - Loaded env from /home/tmi/code/chatbot-ui/.env.local
event - compiled client and server successfully in 2.2s (273 modules)
wait - compiling...
event - compiled successfully in 214 ms (233 modules)
wait - compiling / (client and server)...
event - compiled client and server successfully in 2.8s (5549 modules)
wait - compiling /api/models (client and server)...
event - compiled successfully in 104 ms (43 modules)
wait - compiling...
event - compiled successfully in 1 ms (43 modules)
wait - compiling...
event - compiled successfully in 2 ms (43 modules)
```

API through Chatbot-UI



API through Chatbot-UI



thank you

Pricing

OpenAI API

- requires OpenAI account
- access to all models
- pay per input/output

GPT-4

Model	Prompt	Completion
8K context	\$0.03 / 1K tokens	\$0.06 / 1K tokens
32K context	\$0.06 / 1K tokens	\$0.12 / 1K tokens

Chat

Model	Usage
gpt-3.5-turbo	\$0.002 / 1K tokens

generating four A4 pages

$$1 \text{ token} = \frac{3}{4} \text{ words}$$

$$2,000 \text{ Words (four A4 pages)} \\ = 2,667 \text{ Tokens}$$

gpt-3.5 completion	~ 0.5 Cents
gpt-4 8k completion	~ 16.0 Cents
gpt-4 32k completion	~ 32.0 Cents

- Taxes, credit card fees, ...
- Limits adjustable in account settings