OpenAI Platform

Tokenizer

Learn about language model tokenization

OpenAI's large language models process text using **tokens**, which are common sequences of characters found in a set of text. The models learn to understand the statistical relationships between these tokens, and excel at producing the next token in a sequence of tokens. Learn more.

You can use the tool below to understand how a piece of text might be tokenized by a language model, and the total count of tokens in that piece of text.

| GPT-4o & GPT-4o mini | **GPT-3.5 & GPT-4** | GPT-3 (Legacy) |

You are a responsible financial AI advisor advising a new client with no crypto experience. Based on the following crypto market data: {"coins": [{"id": "bitcoin", "symbol": "btc", "name": "Bitcoin", "current_price": 105419}, {"id": "ethereum", "symbol": "eth", "name": "Ethereum", "current_price": 2555.76}, {"id": "tether", "symbol": "usdt", "name": "Tether", "current_price": 1.0}, {"id": "ripple", "symbol": "xrp", "name": "XRP", "current_price": 2.14}, {"id": "binancecoin", "symbol": "bnb", "name": "BNB", "current_price": 651.89} 1. Which of these coins appears to be the best investment based on strong financial fundamentals? Provide relatable analogies that avoid other traditional investment options like stocks 2. Are there any that resemble meme coins or should be approached with caution? If none present, still mention any currently popular as of today meme and specifically give opinion on trump coin. 3. What red flags should someone watch for if approached in-person to invest in a new

Clear   Show example

Tokens   **Characters**
270   1040

You are a responsible financial AI advisor advising a new client with no crypto experience. Based on the following crypto market data: {"coins": [{"id": "bitcoin", "symbol": "btc", "name": "Bitcoin", "current_price": 105419}, {"id": "ethereum", "symbol": "eth", "name": "Ethereum", "current_price": 2555.76}, {"id": "tether", "symbol": "usdt", "name": "Tether", "current_price": 1.0}, {"id": "ripple", "symbol": "xrp", "name": "XRP", "current_price": 2.14}, {"id": "binancecoin", "symbol": "bnb", "name": "BNB", "current_price": 651.89} 1. Which of these coins appears to be the best investment based on strong financial fundamentals? Provide relatable analogies that avoid other traditional

investment options like stocks 2. Are there any that resemble meme coins or should be approached with caution? If none present, still mention any currently popular as of today meme and specifically give opinion on trump coin . 3. What red flags should someone watch for if approached in-person to invest in a new coin? Respond clearly and concisely.

Text   Token IDs

A helpful rule of thumb is that one token generally corresponds to ~4 characters of text for common English text. This translates to roughly ¾ of a word (so 100 tokens ~= 75 words).

If you need a programmatic interface for tokenizing text, check out our tiktoken package for Python. For JavaScript, the community-supported @dbdq/tiktoken package works with most GPT models.