

# token

## 定义:

Our models understand and process text by breaking it down into tokens. Tokens can be words or just chunks of characters. For example, the word “hamburger” gets broken up into the tokens “ham”, “bur” and “ger”, while a short and common word like “pear” is a single token. Many tokens start with a whitespace, for example “ hello” and “ bye”.

**#important** The number of tokens processed in a given API request depends on the length of *both* your inputs and outputs.

As a rough rule of thumb, 1 token is approximately 4 characters or 0.75 words for English text. One limitation to keep in mind is that your text prompt and generated completion combined must be no more than the model's maximum context length (for most models this is 2048 tokens, or about 1500 words). Check out our [tokenizer tool](#) to learn more about how text translates to tokens.

OpenAI 官网提供了一个文本与 token 长度的计算[工具](#)。

每个 GPT 模型都有一个预设的最大 Tokens 数量。例如，GPT-3 每次调用允许处理的最大 Tokens 数量约为 4096。GPT4则允许处理3万多个 Token， 这个数量包括用户输入和 GPT 输出的所有 Tokens。

## 价格