* Tokens

Um token pode ser uma palavra individual, um caractere, ou até mesmo uma parte de uma palavra. Por exemplo, a frase "Hello World", teria os seguintes tokens:



Total de 3 tokens.

Enquanto isso, essa mesma frase em português, seria dividida da seguinte forma:



Total de 5 tokens.

Dependendo do idioma o processo de tokenização divide palavras de formas diferentes.

De curiosidade como um texto se traduz em tokens: https://platform.openai.com/tokenizer