

Token: O que são?

Vimos em aula o que é o processo de tokenização e como aplicá-lo em uma string.

Em NLP, o que são os tokens gerados neste processo?

Selecione uma alternativa

- A** Token é uma sequência de caracteres separados por um limitador. Esse limitador é sempre o espaço em branco.
- B** Token é uma sequência de caracteres, separados por um limitador, que pode ser um espaço em branco, pontuação ou quebra de linhas, por exemplo.
- C** Tokens são palavras advindas do processo de tokenização de uma frase.