

Token: separando uma frase.



15%

ATIVIDADES
3 de 10DISCORD
ALURAFÓRUM DO
CURSOVOLTAR
PARA
DASHBOARDMODO
NOTURNO

16.0k xp

a

Nesta aula discutimos sobre o processo de tokenizar uma frase, ou seja, dividi-la em pequenos pedaços. Vimos que há algumas formas de realizar esse processo, uma delas com o `split()` e a outra utilizando os métodos `tokenize` da biblioteca NLTK. Levando em consideração o trecho de código abaixo, responda:

```
import nltk
texto_exemplo = "Olá, tudo bem?"
palavras_separadas = nltk.tokenize.word_tokenize(texto_exemplo)
print(palavras_separadas)
```

COPIAR CÓDIGO



Qual alternativa apresenta corretamente a saída da função `print`?

A

['Olá', 'tudo', 'bem']



['Olá', ',', 'tudo', 'bem', '?']



Parabéns, isso mesmo! Quando usamos o `word_tokenize()` não separamos os tokens levando em conta apenas os espaços, como no `split()`. É considerada também a separação entre as pontuações.



15%

C

['Olá,', 'tudo', 'bem?']



PRÓXIMA ATIVIDADE

ATIVIDADES

3 de 10

DISCORD
ALURA

FÓRUM DO
CURSO

VOLTAR
PARA
DASHBOARD

MODO
NOTURNO



16.0k xp

a