

Tradução e Reconhecimento de fala

Usando a biblioteca Googletrans

Tradução

Googletrans

Googletrans é uma biblioteca *Python* gratuita e ilimitada que implementa a *Google Translate API*. Esta biblioteca usa o *Google Translate Ajax API* para fazer chamadas para métodos como detectar e traduzir.

Tradução

Googletrans

Recursos:

- Rápida e confiável – Utiliza os mesmos servidores que o `translate.google.com`
- Detecção automática de idioma
- Traduções em massa
- URL de serviço personalizável
- Pool de conexões (a vantagem de usar `requests.Session`)
- Suporte a HTTP/2

Tradução

Googletrans

Nota sobre o uso da biblioteca:

- O limite máximo de caracteres em um único texto é 15k.
- Devido às limitações da versão da *web* do *google translate*, essa *API* não garante que a biblioteca funcione corretamente em todos os momentos. (então, por favor, use esta biblioteca se você não se importa com a estabilidade.)
- Se você quiser usar uma *API* estável, é altamente recomendável usar a *API* de tradução oficial do *Google*:

(<https://cloud.google.com/translate/docs/>).

- Se você receber um erro *HTTP 5xx* ou erros como *#6*, é provável que o *Google* tenha banido seu endereço *IP*.

Tradução

Googletrans

Podemos instalar com o *Pypi*:

```
$ pip install googletrans
```

No momento em que gravo esta aula existe um problema com a biblioteca, porém, a solução já foi reportada no projeto que está hospedado no *Github*, mas como o *pull request* ainda não foi aceito e a biblioteca continua com o problema, vou desinstalar a versão oficial e instalar uma versão não oficial que foi corrigida:

```
$ pip uninstall googletrans
```

```
$ git clone https://github.com/BoseCorp/py-googletrans.git
```

```
$ cd ./py-googletrans
```

```
$ python setup.py install
```

Tradução

Googletrans

Veja um exemplo básico de uso:

```
from googletrans import Translator

translator = Translator()
# Se não for informada a linguagem de origem, o google translate tenta detectar a linguagem.
# A linguagem destino será inglês por padrão
print(translator.translate('안녕하세요.'))
# Aqui estamos definindo que a linguagem destino será Japonês
print(translator.translate('안녕하세요.', dest='ja'))
# Aqui estamos definindo a linguagem origem igual a Latin e a destino igual a Português
print(translator.translate('veritas lux mea', src='la', dest='pt'))
```

```
Translated(src=ko, dest=en, text=Hi., pronunciation=None, extra_data="{ 'translat..."
```

```
Translated(src=ko, dest=ja, text=こんにちは。 , pronunciation=Kon'nichiwa., extra_data="{ 'translat..."
```

```
Translated(src=la, dest=pt, text=A verdade é minha luz, pronunciation=veritas lux mea, extra_data="{ 'translat..."
```

Tradução

Googletrans

Podemos definir o domínio que será utilizado para tradução, se fornecermos vários endereços, ele irá escolher aleatoriamente.

```
from googletrans import Translator

translator = Translator(service_urls=[
    'translate.google.com',
    'translate.google.com.br',
])

print(translator.translate('veritas lux mea', src='la', dest='pt'))
```

Tradução

Googletrans

Podemos traduzir uma lista de strings (Tradução em massa).

```
from googletrans import Translator
translator = Translator()
lista = ['Olá, tudo bem?', 'Bom dia', 'Eu tenho um cachorro.']

translations = translator.translate(lista, dest='ko')
for translation in translations:
    print(translation.origin, ' -> ', translation.text)
```

```
Olá, tudo bem? -> 안녕, 안녕하세요?
Bom dia -> 좋은 아침.
Eu tenho um cachorro. -> 나는 개가있다.
```


Tradução

Googletrans

Detectando o idioma.

```
from googletrans import Translator

translator = Translator()
print(translator.detect('이 문장은 한글로 쓰여졌습니다. '))
print(translator.detect('この文章は日本語で書かれました. '))
print(translator.detect('This sentence is written in English. '))
print(translator.detect('Tiu frazo estas skribita en Esperanto. '))
print(translator.detect('Esta frase está escrita em Português'))
```

```
Detected(lang=ko, confidence=1)
Detected(lang=ja, confidence=1)
Detected(lang=en, confidence=1)
Detected(lang=eo, confidence=1)
Detected(lang=pt, confidence=1)
```

Tradução

Googletrans

translate.google.com usa um *token* para autorizar as solicitações. Desta forma o *Google* cobra pelo uso de sua *API*. GoogleTrans tem uma classe interna chamada *TokenAcquirer* para gerar um *token* válido para nossa aplicação. Veja esta classe sendo utilizada.

```
from googletrans.gtoken import TokenAcquirer

acquirer = TokenAcquirer()
texto = 'Olá!'
tk = acquirer.do(texto)
print(tk)
```

Tradução

Googletrans

Traduzindo um texto informado pelo usuário (Inglês/Português).

```
from googletrans import Translator

texto_en = input("Write a sentence: ")
translator = Translator()
texto_pt = translator.translate(texto_en, src="en", dest="pt")
print("Original text:", texto_en)
print("Translation:", texto_pt.text)
```

Tradução

Googletrans

Traduzindo um texto informado pelo usuário
(Português/Inglês).

```
from googletrans import Translator

texto_pt = input("Escreva uma frase: ")
translator = Translator()
texto_en = translator.translate(texto_pt, src="pt", dest="en")
print("Texto original:", texto_pt)
print("Tradução:", texto_en.text)
```

FIM