

Processamento de Linguagem Natural com spaCy

Jones Granatyr



Processamento de linguagem natural

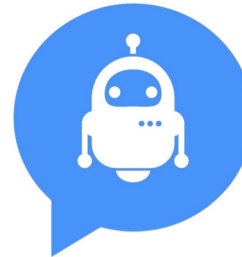
- Computadores que analisam e entendem a linguagem humana



Speech Transcription



Neural Machine
Translation (NMT)



Chatbots



Q&A



Text Summarization



Image Captioning



Video Captioning

Conteúdo

Bill Gates PER

nasceu em

Seattle LOC

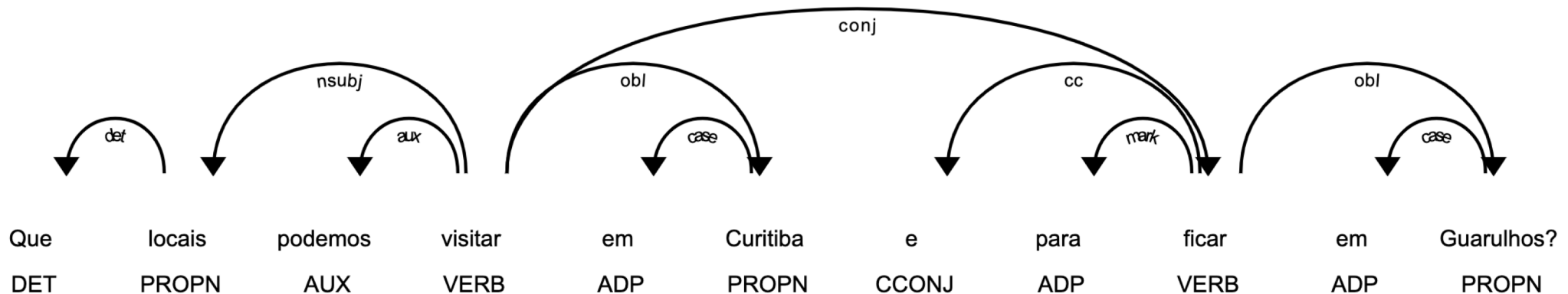
em 28/10/1955 e foi o criador da

Microsoft ORG

- Marcação POS (part-of-speech)
- Lematização x stemização
- Reconhecimento de entidades nomeadas
- Parsing de dependências
- Semelhanças entre palavras
- Classificação de sentimentos – frases em português de alegria e medo
- *Que locais podemos visitar em Curitiba e para ficar em Guarulhos?*
 - *Curitiba para visitar*
 - *Guarulhos para ficar*

```
doc = pln('encontrei encontraram encontrarão encontrariam cursando curso cursei')  
[token.lemma_ for token in doc]
```

```
['encontrar',  
'encontrar',  
'encontrar',  
'encontrar',  
'cursar',  
'cursar',  
'cursar']
```



spaCy

- PLN avançado escrita em Python e Cython
- Construída por Matthew Honnibal
- Considerado o parser sintático mais rápido do mundo
- Possui internamente modelos de redes neurais em vários idiomas
- Uso em larga escala (auxiliado com o Cython)
- Processamento de textos para deep learning

Pré-requisitos

- Lógica de programação, principalmente estruturas condicionais e de repetição
- Básico da linguagem Python (desejável)
- Não são necessários conhecimentos prévios sobre inteligência artificial
- Nível do curso: iniciante