

**Mercado de  
Trabalho**

**Mestrado**

**Linguística  
e NLP**

**Por que  
Estatística?**

**Beatriz Albiero**

**[www.linkedin.com/in/  
beatriz-albiero/](https://www.linkedin.com/in/beatriz-albiero/)**



**Por que Estatística?**

# Por que Estatística?

- Profissão do futuro

# Por que Estatística?

- Profissão do futuro
- Salário

# Por que Estatística?

- Profissão do futuro
- Salário
- Flexibilidade no Mercado

# Por que Estatística?

- Profissão do futuro
- Salário
- Flexibilidade no Mercado
- Demanda

# Linguística

- Linguística Histórica

# Linguística

- Linguística Histórica
- Fonética e Fonologia



# Linguística

- Linguística Histórica
- Fonética e Fonologia
- Sintaxe

# Linguística

- Linguística Histórica
- Fonética e Fonologia
- Sintaxe
- Semântica

# Linguística

- Linguística Histórica
- Fonética e Fonologia
- Sintaxe
- Semântica
- Pragmática

# Linguística

- Aquisição de Linguagem
- Neurolinguística
- Psicolinguística
- Linguística Computacional

# Linguística Computacional

Campo interdisciplinar preocupado com os modelos de linguagem natural (com propostas estatísticas ou baseadas em regras) de um ponto de vista computacional, bem como o estudo de abordagens computacionais adequadas para questões linguísticas.

# **PLN - Processamento de Linguagem Natural:**

**O PLN visa tornar a relação humano-computador o  
mais natural possível.**

# PLN - Processamento de Linguagem Natural:

O PLN visa tornar a relação humano-computador o  
mais natural possível.



# PLN – Questões:

- Conversão áudio-texto
- Predição de palavras e sentenças
- Análise de Sentimento
- Tradução automática



# PLN – Questões:

## Predição de palavras e sentenças

The 20 funniest suggestions from Google Suggest



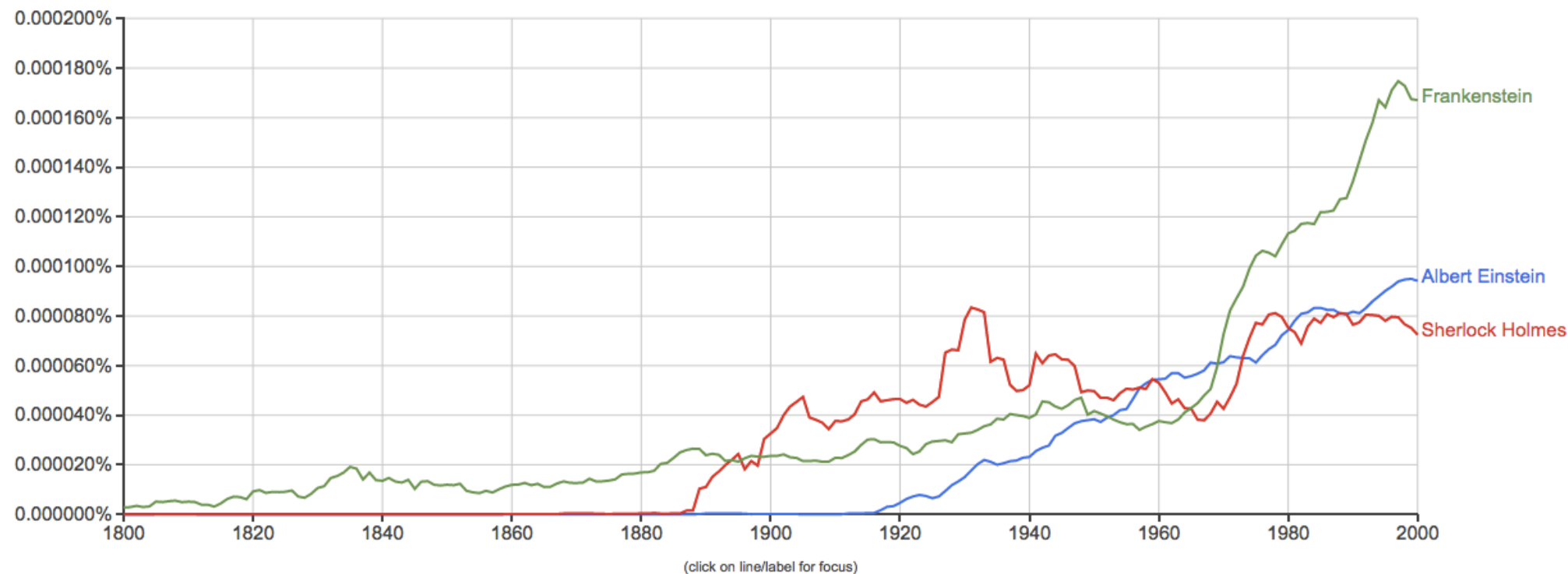
what would a	<a href="#">Advanced Search</a>
	<a href="#">Tools</a>
what would audrey do	5,710,000 results
what would an accountant use a spreadsheet for	520,000 results
what would a banker use a spreadsheet for	1,180,000 results
what would aliens look like	9,940,000 results
what would a scientist use a spreadsheet for	609,000 results
what would a teacher use a spreadsheet for	705,000 results
what would a shopkeeper use a spreadsheet for	18,200 results
what would an engineer use a breadboard for	48,800 results
what would an alien look like	14,100,000 results
what would a chair look like if your knees bent the other way	507,000 results
<a href="#">close</a>	

[Adverti](#)

# PLN – Questões:

## Predição de palavras e sentenças

<https://books.google.com/ngrams>



# PLN – Questões:

*Análise de sentimento*

# PLN – Questões:

## Análise de sentimento

"Eu gosto do Felipe mas ele é um  
babaca."

# PLN – Questões:

## Análise de sentimento

"Eu gosto do Felipe mas ele é um babaca."



# PLN – Questões:

## Análise de sentimento

"Eu gosto do Felipe mas ele é um babaca."

"Eu adoro essa marca mas esse produto é uma porcaria."



# PLN – Questões:

Correção automática:

# PLN – Questões:

## Tradução automática:

My loneliness is killing me (and I)  
I must confess I still believe (still  
believe)

When I'm not with you I lose my mind

Give me a sign

Hit me baby one more time



# PLN – Questões:

## Tradução automática:

Magányom megöli (és én)  
Be kell vallanom, még mindig hiszek  
(még mindig hiszek)  
Ha nem vagyok veled, elveszítem az  
eszemet  
Adj egy jelet  
Üss meg bébi még egyszer

hungaro

# PLN – Questões:

## Tradução automática:

Magányom megöli (és én)  
Be kell vallanom, még mindig hiszek  
(még mindig hiszek)  
Ha nem vagyok veled, elveszítem az  
eszemet  
Adj egy jelet  
Üss meg bébi még egyszer

hungaro

My loner kills me (and me)  
I have to confess, I still  
believe (I still believe)  
If I'm not with you, I lose  
my mind  
Give me a mark  
Have a baby again

# PLN – Questões:

Tradução automática:

Corpus  
Sick - PT

# PLN – Questões:

Tradução automática:

"A white dog is standing in a lake."

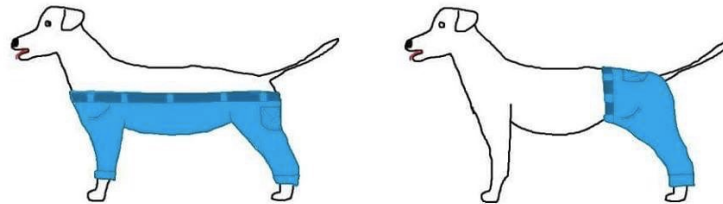
Corpus  
Sick - PT

# PLN – Questões:

Correção automática:

"A white dog is standing in a lake."

If a dog wore pants would he wear them  
like this                      or                      like this?





**PLN – Questões:**

**Natural Language Inference:**

# PLN – Questões:

## Natural Language Inference:

**P:** Um jogo de futebol está sendo jogado por alguns homens.

**H:** Alguns homens estão praticando algum esporte.

**Relação:** Acarretamento

**P:** Alguns homens estão praticando algum esporte.

**H:** Um jogo de futebol está sendo jogado por alguns homens.

**Relação:** Neutro

**P:** Todos os homens estão praticando algum esporte.

**H:** Nenhum homem está praticando esporte.

**Relação:** Contradição.

## + Questões:

- Correção automática
- Chatbots
- Text Mining
- Sumarização
- Classificação



## **+ Sobre Linguística:**

- **"Steven Pinker: Linguistics as a Window to Understanding the Brain" no Youtube**
- **Livro "The Language Instinct" - Steven Pinker**
- **Livro "Speech and Language Processing. an Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition." - JURAFSKY e MARTIN**



**GLiC - USP**

# **GLiC - USP**

- **Marcelo Barra e Marcos Lopes**

# GLiC - USP

- Marcelo Barra e Marcos Lopes
- [glicusp.wordpress.com/](http://glicusp.wordpress.com/)

# GLiC - USP

- Marcelo Barra e Marcos Lopes
- [glicusp.wordpress.com/](http://glicusp.wordpress.com/)
- NLI e Sick-PT
- Sentences Involving  
Compositional Knowledge

# **Meu Projeto**

**Um Estudo Sobre as Irregularidades Verbais  
do Português Brasileiro Via Redes Neurais  
Artificiais**

# Meu Projeto

**Um Estudo Sobre as Irregularidades Verbais  
do Português Brasileiro Via Redes Neurais  
Artificiais**

**[github.com/beatrizalbiero/  
MsResearch](https://github.com/beatrizalbiero/MsResearch)**

# Workshop do GLiC 2017

**Aprendizado de Máquina em  
Linguagem Natural**

[github.com/beatrizalbiero/ling-comp2017](https://github.com/beatrizalbiero/ling-comp2017)

+ Felipe Salvatore

[github.com/felipessalvatore](https://github.com/felipessalvatore)





# **Mercado de Trabalho**



Mercado de Trabalho

TODO MUNDO PRECISA DE  
NLP

# Referências

20 Funniest Suggestions from Google Suggest:

<https://www.telegraph.co.uk/technology/google/6161567/The-20-funniest-suggestions-from-Google-Suggest.html>

Corpus Sick:

<http://clic.cimec.unitn.it/composes/sick.html>

N-Gramas do Google:

<https://books.google.com/ngrams>

"Steven Pinker: Linguistics as a Window to Understanding the Brain":

[https://youtu.be/Q-B\\_ONJIEcE](https://youtu.be/Q-B_ONJIEcE)