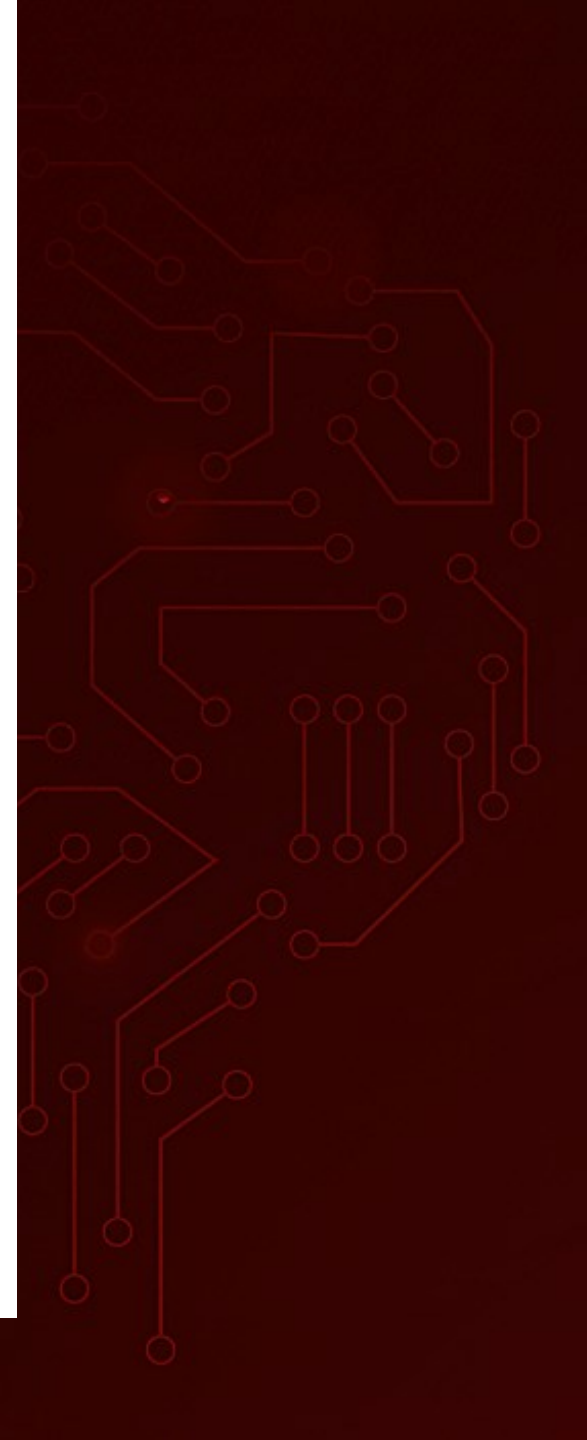


PROCESSAMENTO DE LINGUAGEM NATURAL

Introdução ao Processamento de Linguagem Natural



TÓPICOS

1. Linguagens

2. PLN

3. Níveis Linguísticos

4. Aplicações





LINGUAGENS

LINGUAGEM NATURAL X LINGUAGEM ARTIFICIAL

➤ Linguagem natural

Tem ambiguidade, é usual e informal, está viva e é muito flexível!

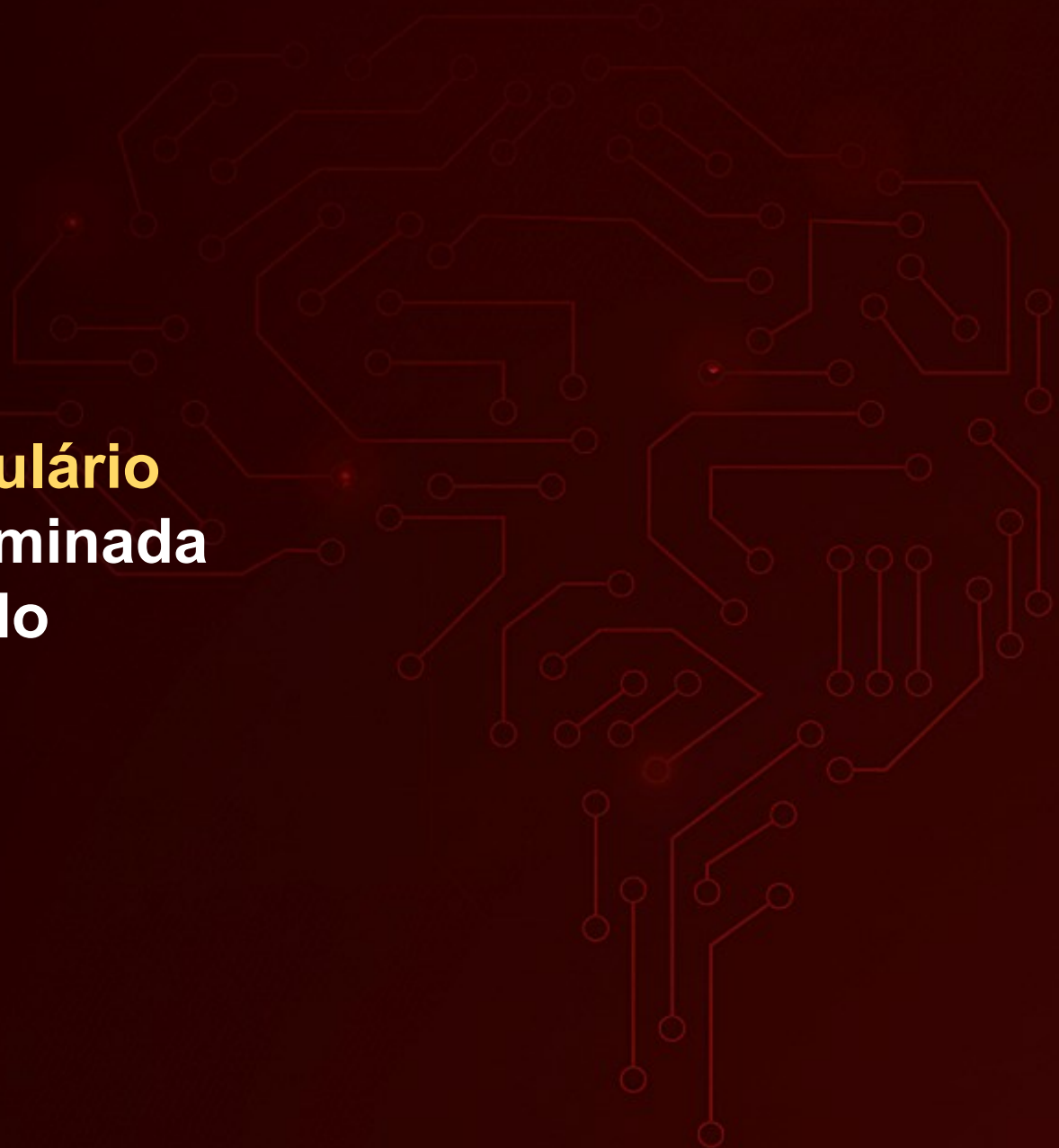
➤ Linguagem artificial

Não tem ambiguidade, é mais formal e poucos conhecem, não admite tanta flexibilidade!

LINGUAGENS

O QUE É A LINGUAGEM?

“Sistema de **símbolos de um vocabulário** que, quando colocados numa determinada **ordem** e expressos num determinado **contexto**, emitem um **significado**.”



PLN

Linguística + Computação

X

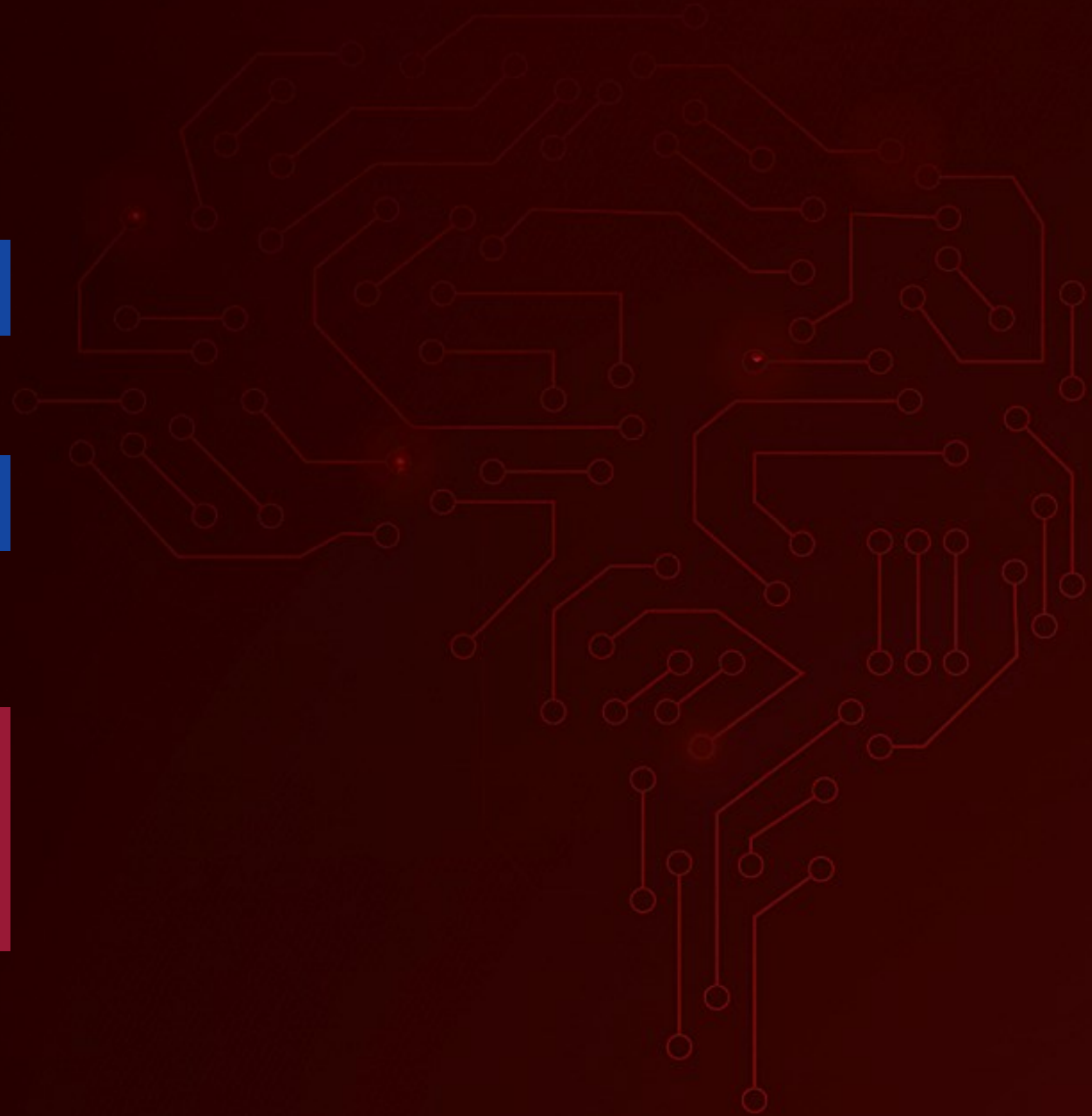
Computação + Linguística

LINGUÍSTICA

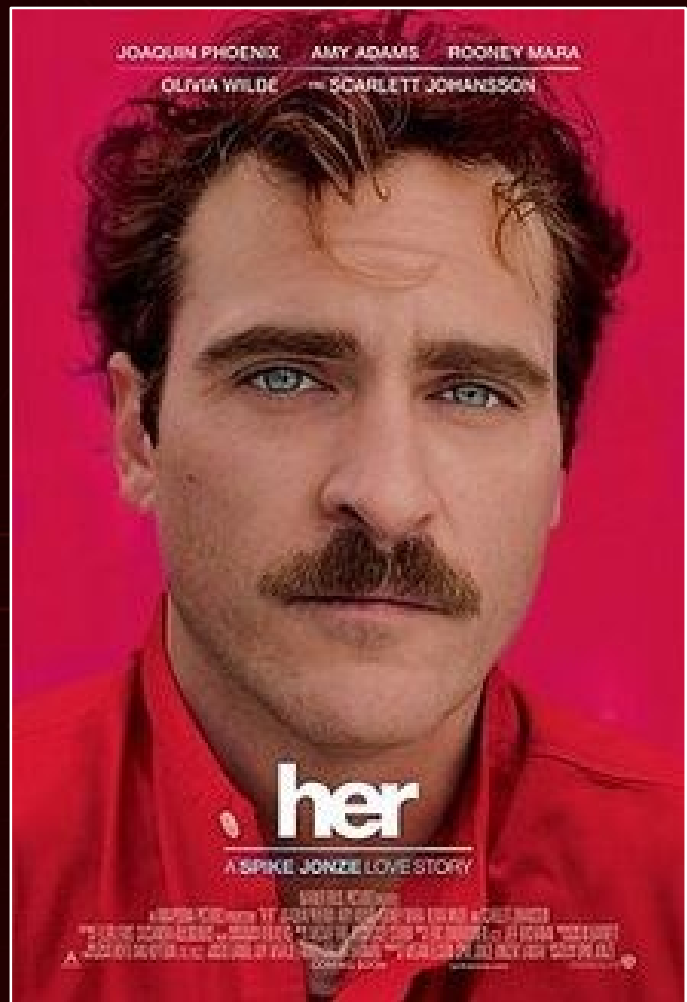
➤ **Modela a
linguagem**

COMPUTAÇÃO

➤ **Implementa
os modelos**



PRA QUE SERVE PLN?

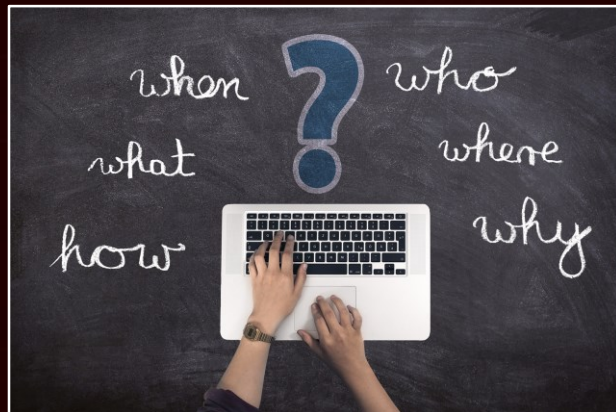


Fonte: <https://pt.wikipedia.org/wiki/Her>

PLN

OBJETIVOS

- Tornar os computadores aptos a processarem a língua natural
- Processar = entender, gerar, extrair conhecimento útil, comunicar-se etc.



PLN

Quão difícil é processar automaticamente a linguagem natural?

O menino foi para a escola de ônibus.

Quão difícil é processar automaticamente a linguagem natural?

O	menino	foi	para	a	escola	de	ônibus	.
---	--------	-----	------	---	--------	----	--------	---

Tokenização
Tokenization

Quão difícil é processar automaticamente a linguagem natural?

O menino foi para a escola de ônibus .

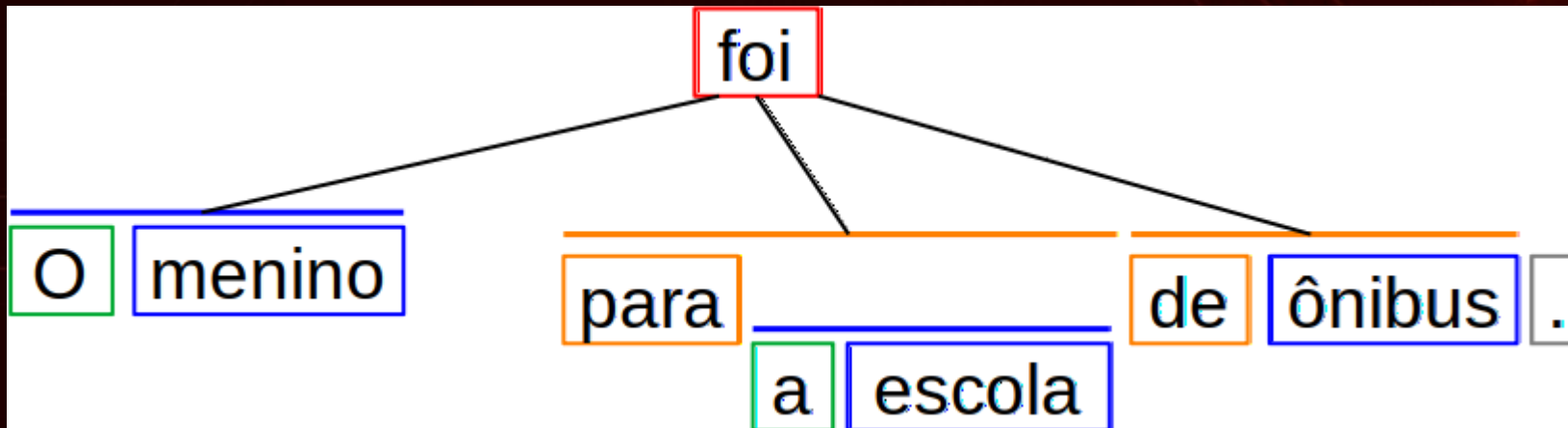
Etiquetação
morfofossintática
Part-of-speech tagging

Quão difícil é processar automaticamente a linguagem natural?

O menino foi para a escola de ônibus .

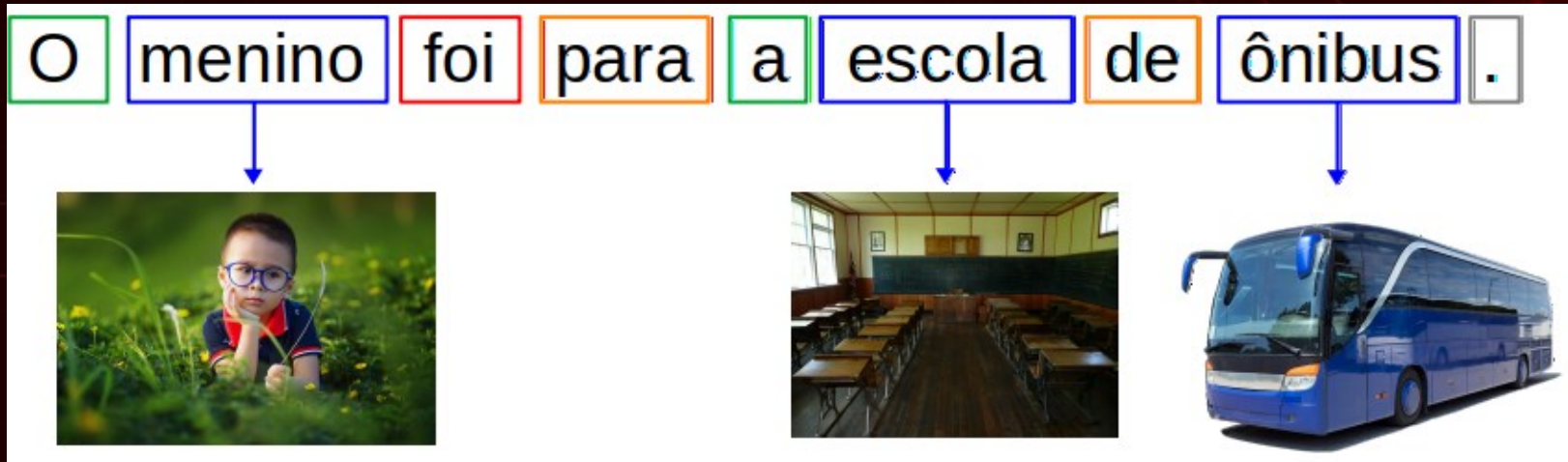
Identificação de sintagmas
Chunking

Quão difícil é processar automaticamente a linguagem natural?



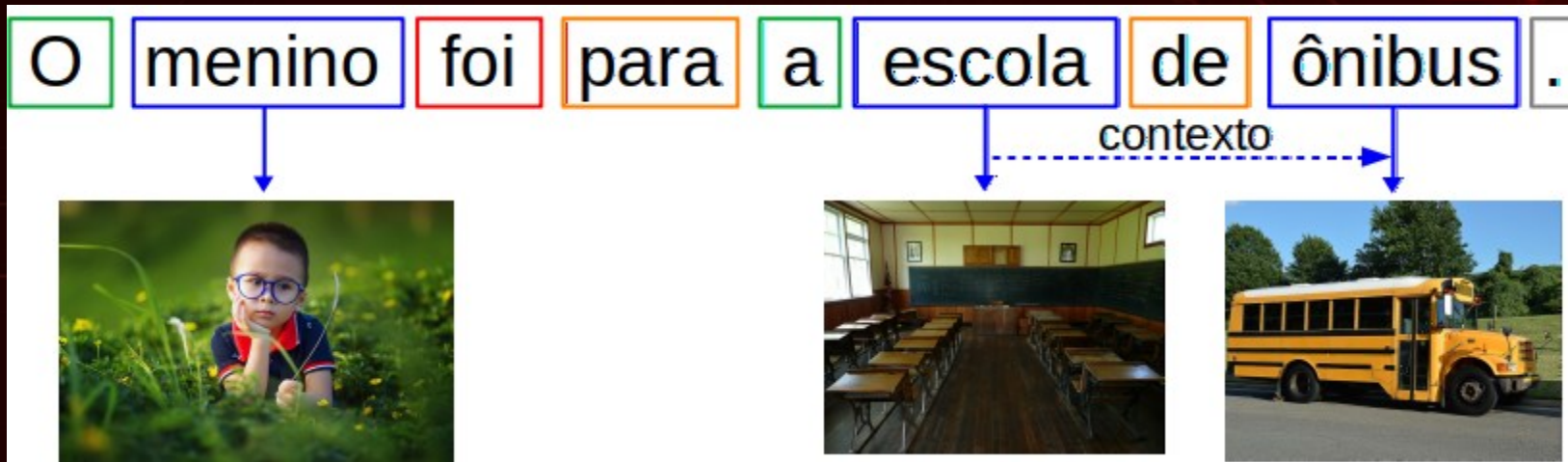
Análise sintática
Parsing

Quão difícil é processar automaticamente a linguagem natural?



Análise Semântica
Semântica Lexical

Quão difícil é processar automaticamente a linguagem natural?



Análise de contexto
Pragmática

PLN

Quão difícil é processar automaticamente a linguagem natural?

O menino foi para a escola de ônibus .



*Grounded meaning
Natural Language Understanding*

NÍVEIS LINGUÍSTICOS

Fonética e Fonologia

- Estudo dos sons das línguas

Morfologia

- Estudo das unidades linguísticas

Sintaxe

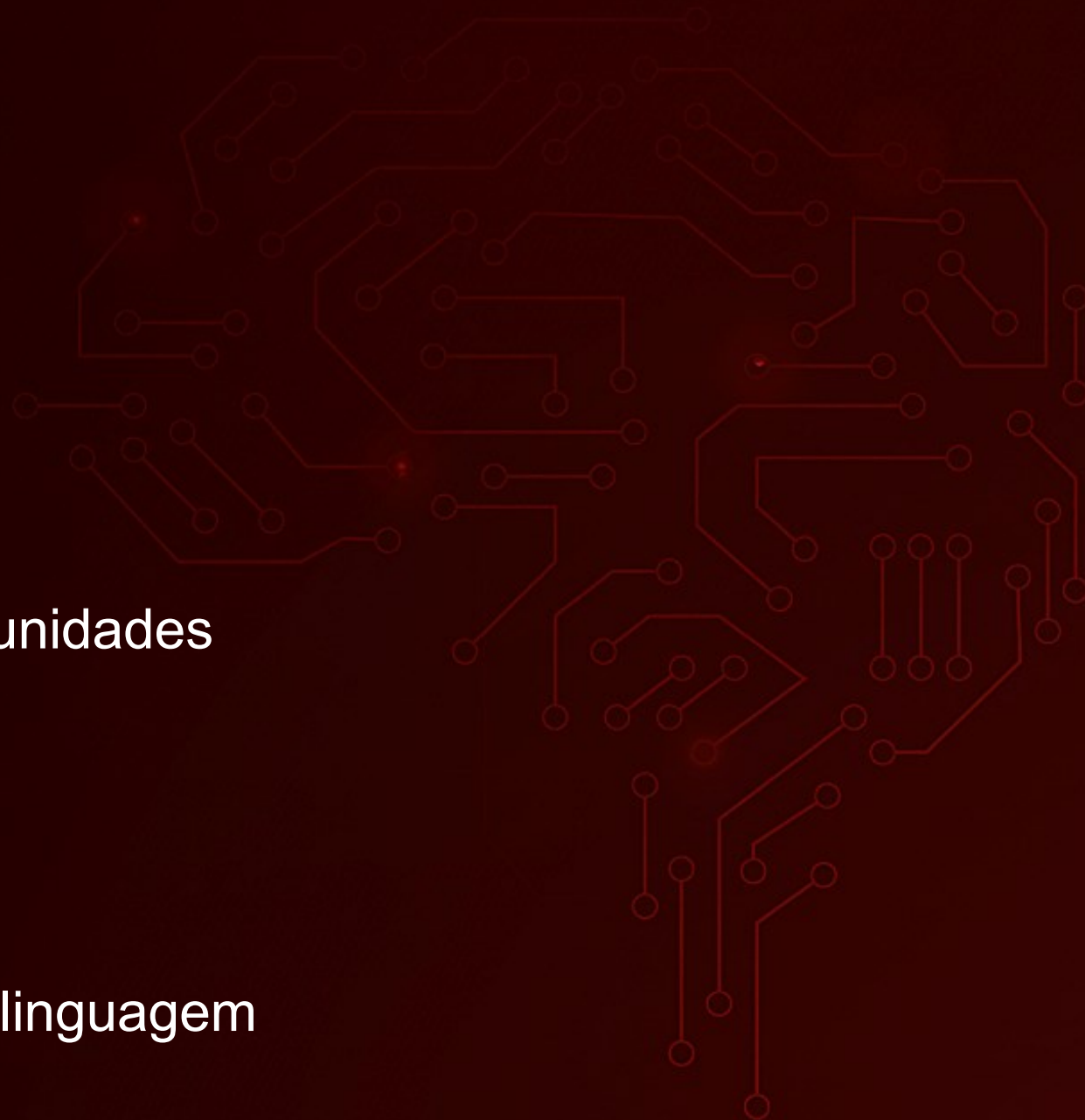
- Estudo das regras que ordenam as unidades linguísticas

Semântica

- Estudo dos significados

Pragmática

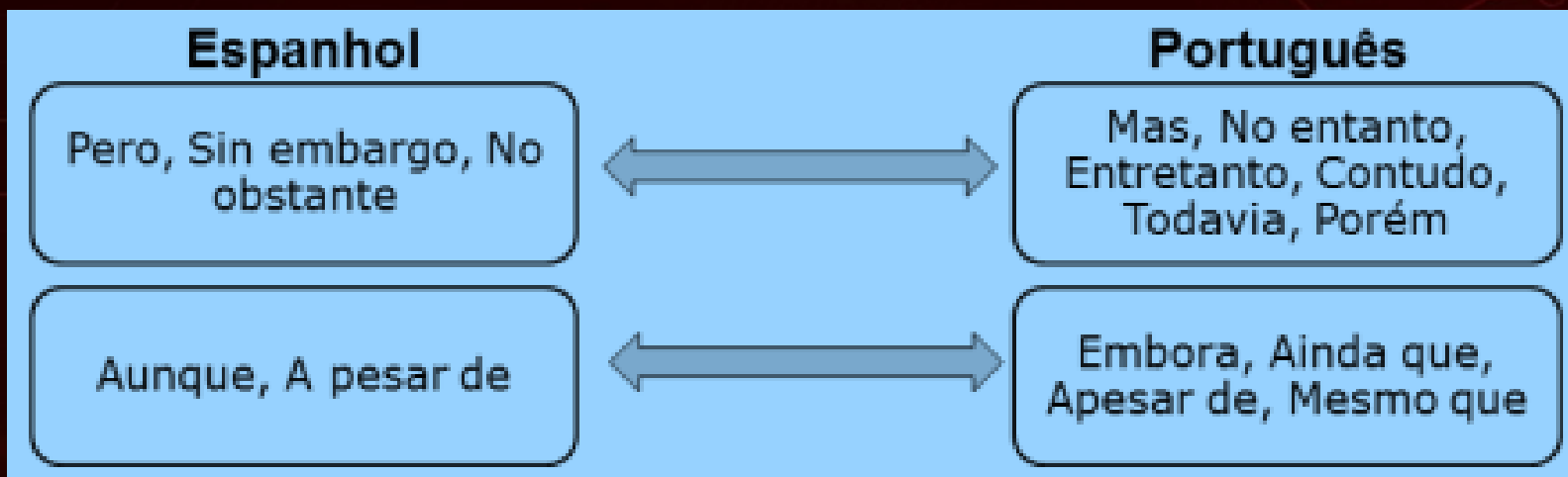
- Estudo da influência do contexto na linguagem



ESTRATÉGIAS

LINGUÍSTICA (SIMBÓLICA)

- O conhecimento é explicitamente mapeado por especialistas para ser utilizado pela máquina



Fonte: TORRES et al. (2014)

ESTRATÉGIAS

ESTATÍSTICA

- A melhor saída é determinada por medidas estatísticas

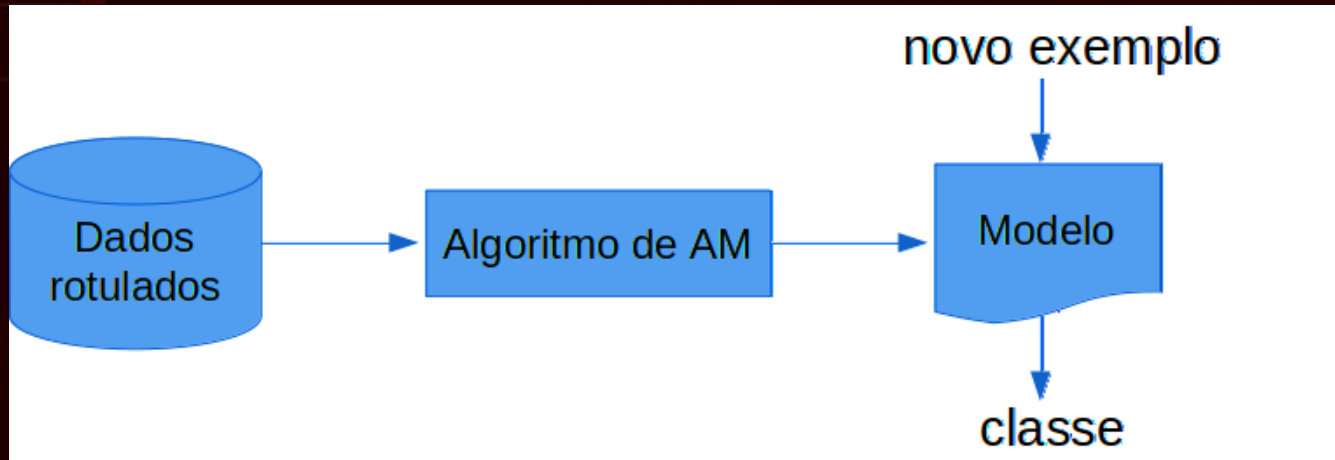
Maria	no	daba	una	bofetada	a	la	bruja	verde
Mary	not	give	a	slap	to	the	witch	green
	did not		a slap	by			green witch	
	no		slap		to the			
	did not give				to			
					the			
			slap			the witch		

Fonte: <http://www.statmt.org/moses/?n=moses.background>

ESTRATÉGIAS

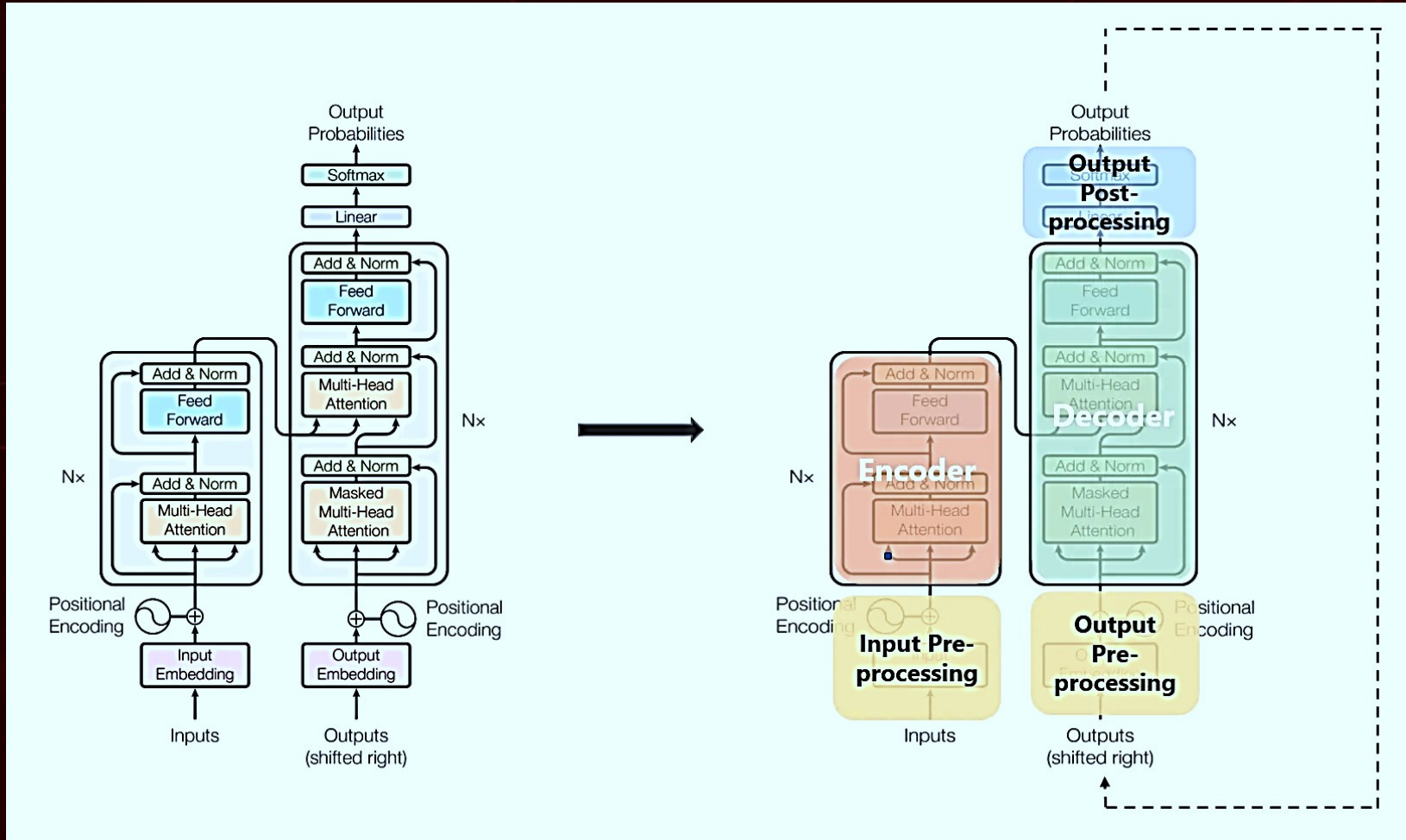
APRENDIZADO DE MÁQUINA TRADICIONAL (*SHALLOW LEARNING*)

- Aprendizado supervisionado
- Anotação do conjunto de treinamento
- Classificação



ESTRATÉGIAS

APRENDIZADO PROFUNDO (*DEEP LEARNING*) – TRANSFORMERS



APLICAÇÕES

Análise de Sentimentos



Chatbot



Assistente de voz



Tradução automática



Detecção de *fake news*



O QUE VIMOS?

- **Linguagens**
- **PLN**
- **Níveis Linguísticos**
- **Aplicações**



PRÓXIMA VIDEOAULA

- Sequências de caracteres, *tokens* e palavras

REFERÊNCIAS

- **Curso de Processamento de Linguagem Natural**
 - **Profa. Helena Caseli (UFSCar)**
- **Curso de Processamento de Linguagem Natural**
 - **Prof. Thiago Pardo (ICMC-USP)**
- **Curso de Linguística Computacional**
 - **Prof. Thiago Castro Ferreira (UFMG)**