A large orange shape on the left side of the slide, consisting of a rectangle on the left and a quarter-circle on the right.

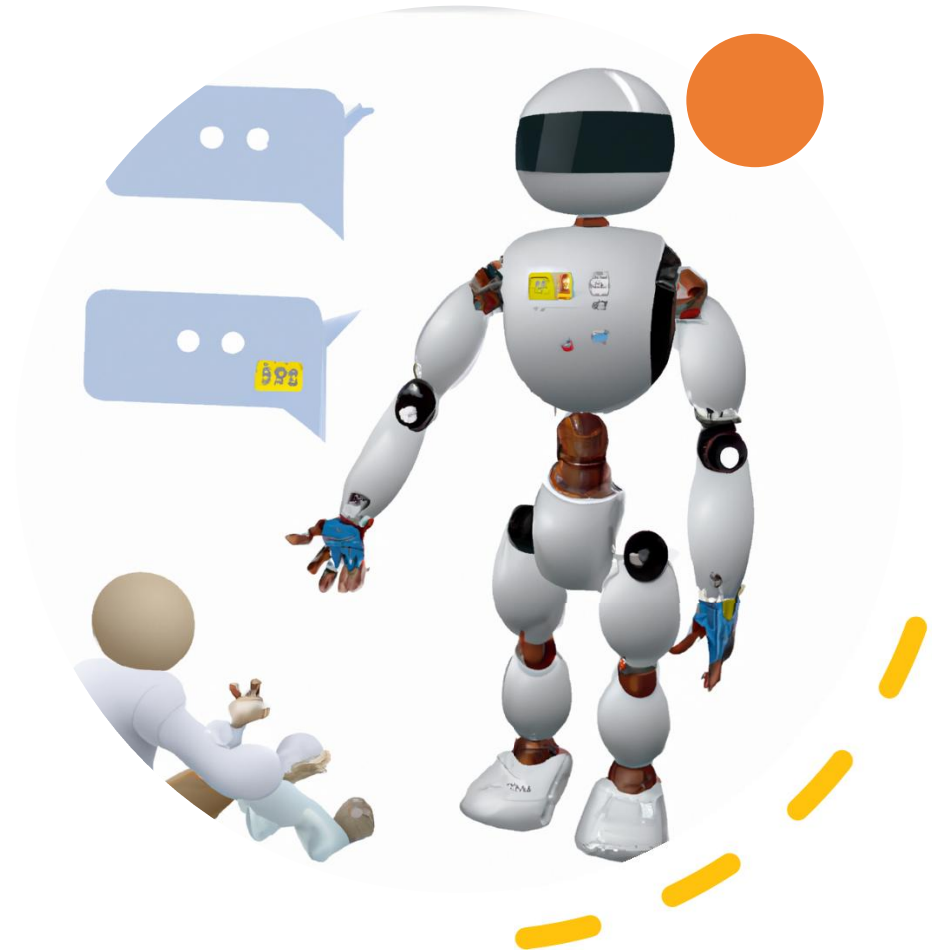
“Grandes” modelos de NLP pré- treinados

- BERT: Google
- XLNet: Carnegie Mellon University e Google
- RoBERTa: Facebook
- T5: Google
- ERNIE: Baidu
- GPT: OpenAI



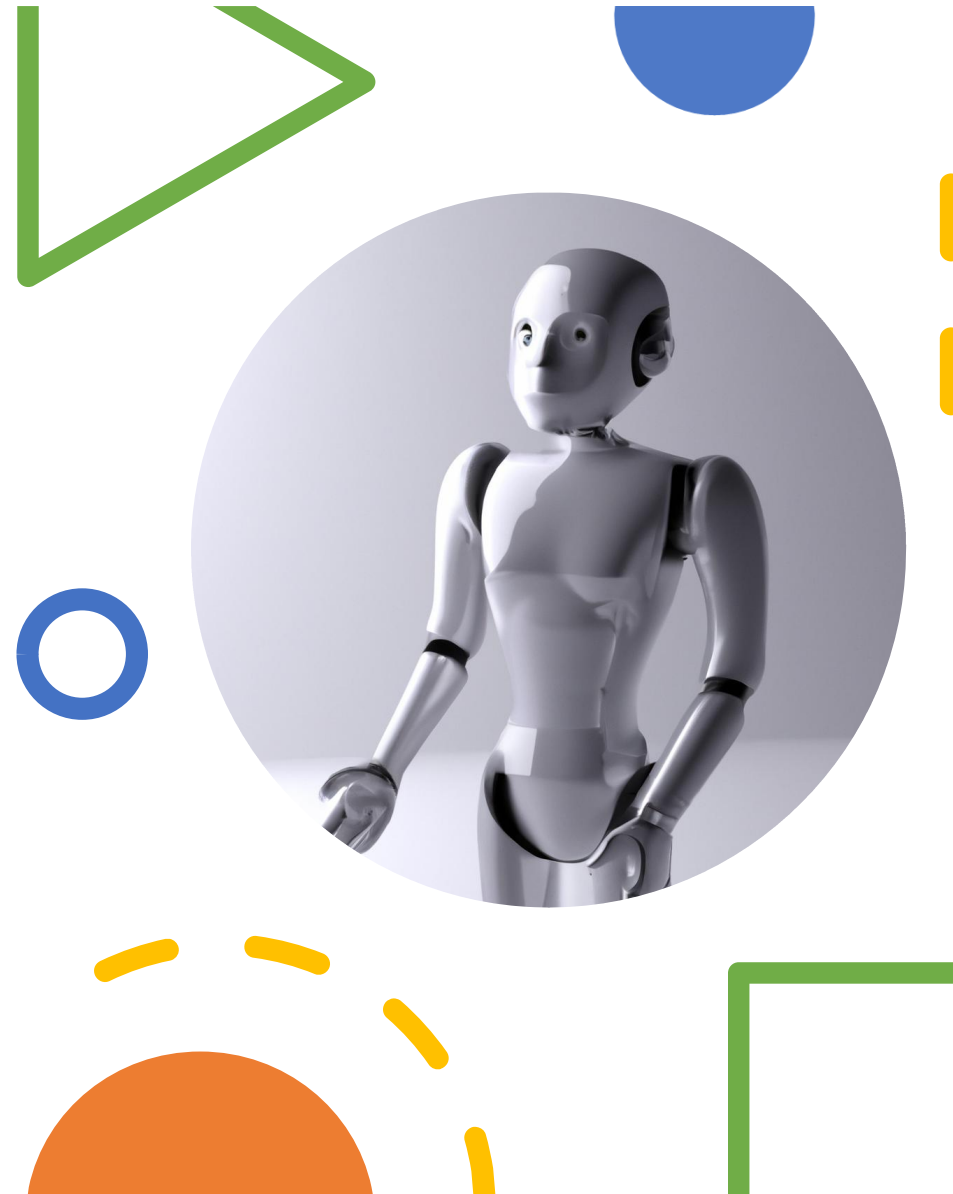
GPT: O que é?

- Modelos de Inteligência artificial para NLP, baseados em Transformers.



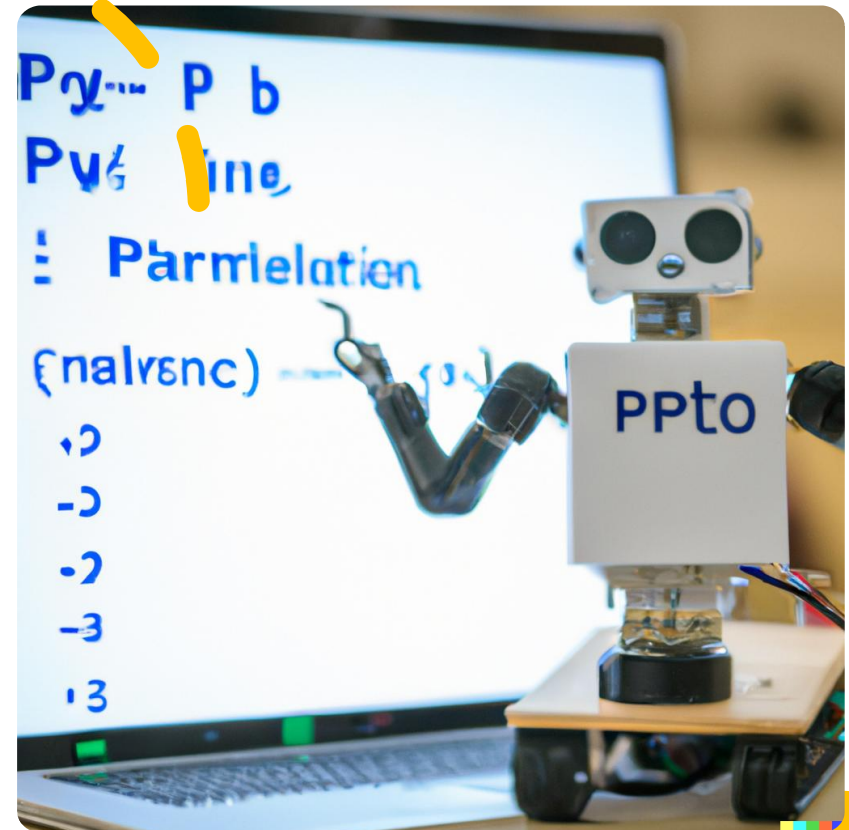
GPT: Significado

- **Generative** : capacidade de produzir novos textos a partir de padrões e estruturas dos dados de treinamento
- **Pre-training**: pré-treinado com texto gerado por humanos
- **Transformer**: baseado na arquitetura de Transformers



GPT: Versões

- GPT foi lançado em 2018, 1,5 bilhões de parâmetros
- GPT-2 criado em 2019. 175 bilhões de parâmetros
- GPT-3 criado em 2020. 1.5 trilhões de parâmetros
- Cada modelo foi treinado em um conjunto maior e mais diversificado de texto gerado por humanos.



GPT: Parâmetros

- Pesos das RNA
- Mecanismos de Atenção
- Controle de Geração de Texto



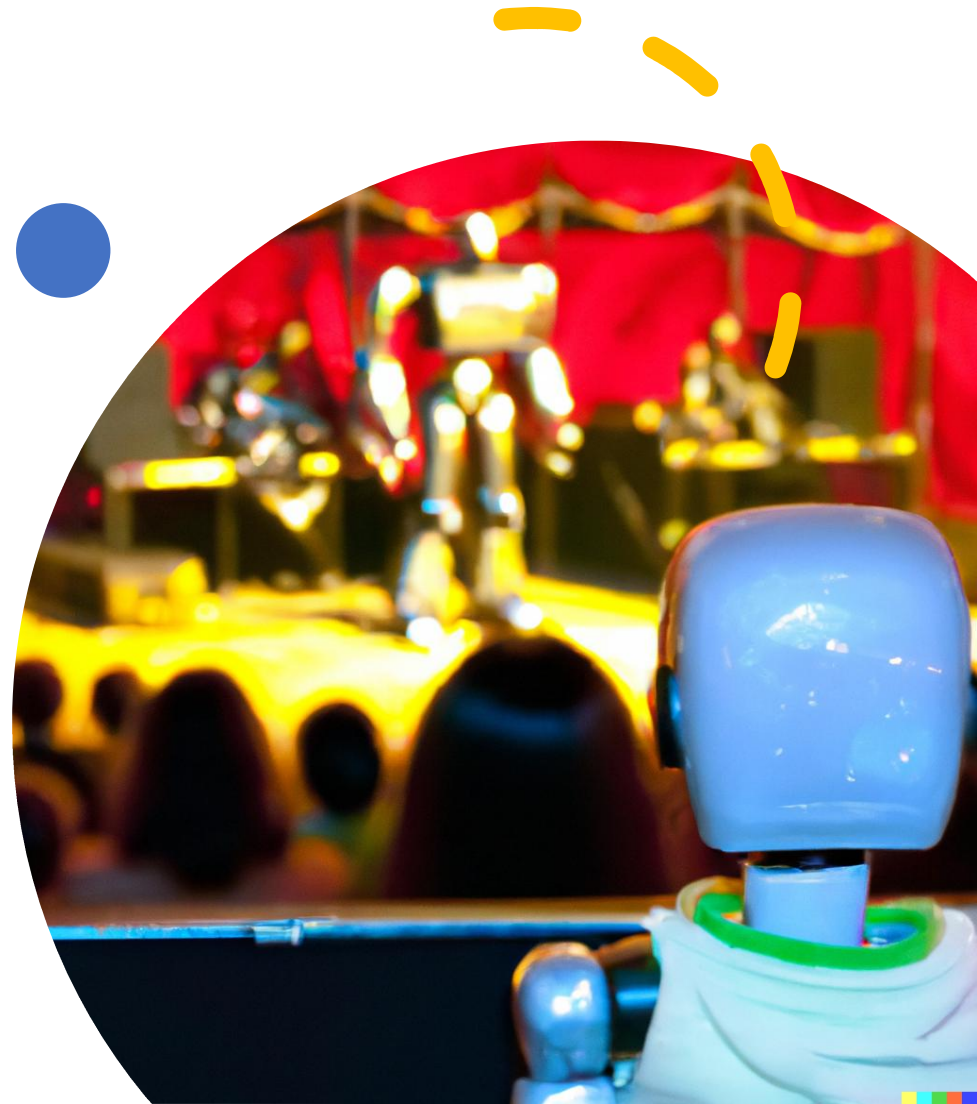
GPT: Treinamento

- **Pre-training:** treino não supervisionado em grandes volumes de texto. Utilizando técnicas como “masked language modeling” e “next-sentence prediction”
- **Fine-tuning:** ajuste para tarefas específicas, como tradução, resumo etc.



GPT: Idiomas

- Inglês
- Frances
- Espanhol
- Alemão
- Italiano
- Chinese
- Portugues
- Etc



Em geral,
modelos de
NLP são
melhores em
inglês!

