

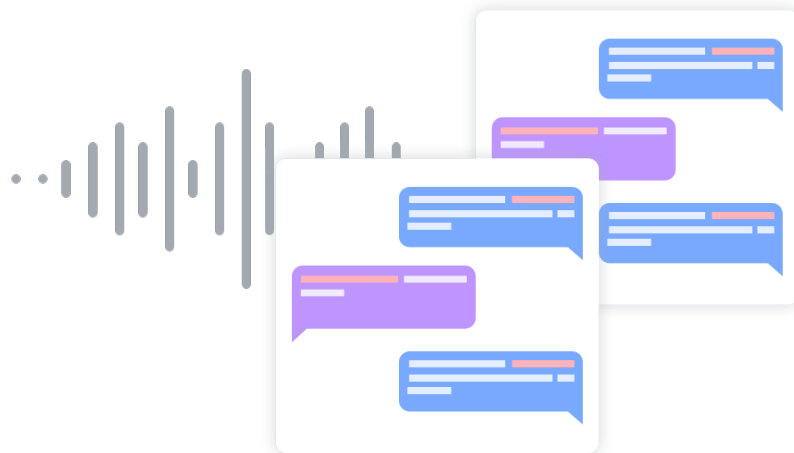
# ASSISTENTES VIRTUAIS EM PYTHON: O GUIA PARA INICIANTES



# CONTEÚDO DO CURSO

- Visão geral sobre reconhecimento da fala e sintetização da voz
- Implementação

# VISÃO GERAL: RECONHECIMENTO DA FALA

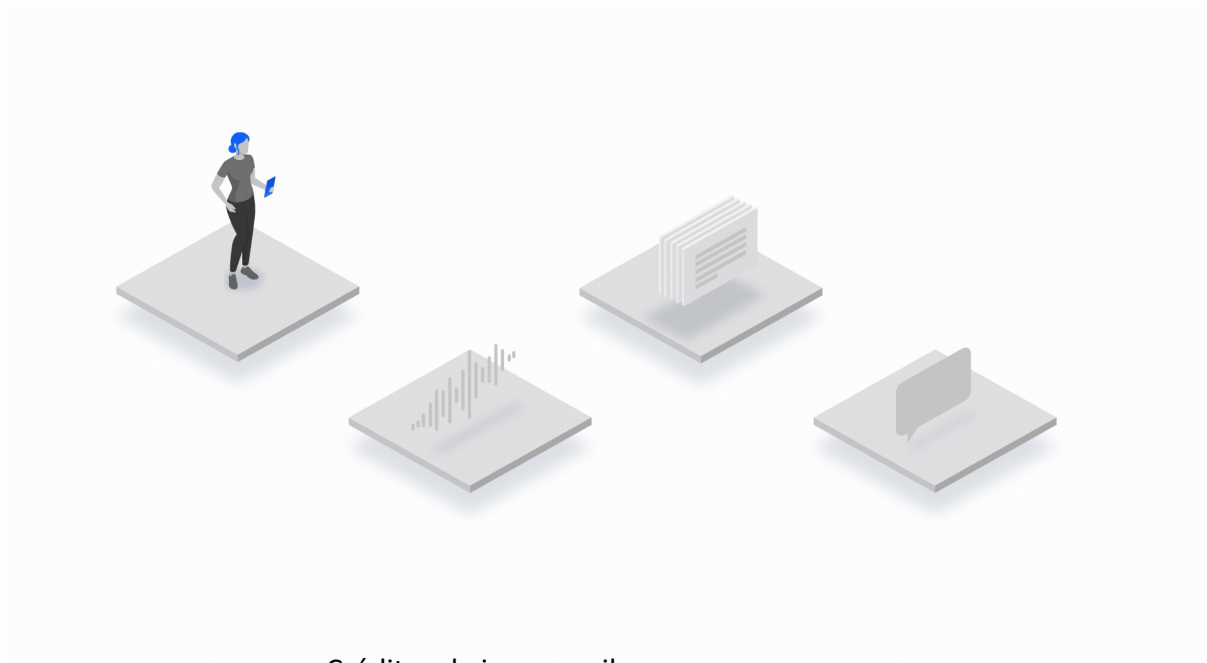


Créditos da imagem: ibm.com

- Também encontramos outros nomes para este mesmo campo, como *automatic speech recognition* (reconhecimento automático de fala), *computer speech recognition* (reconhecimento de fala por computador) ou STT, de *speech to text* (fala para texto).

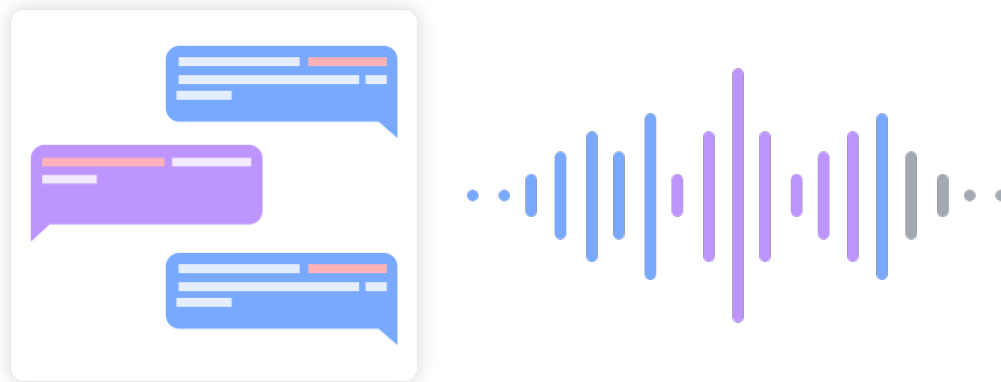
# VISÃO GERAL: RECONHECIMENTO DA FALA

- O algoritmo analisa estatisticamente os fonemas capturados e os compara a uma grande biblioteca de palavras, frases e sentenças conhecidas. O programa determina o que o usuário provavelmente estava dizendo e o envia como texto ou emite um comando do computador.



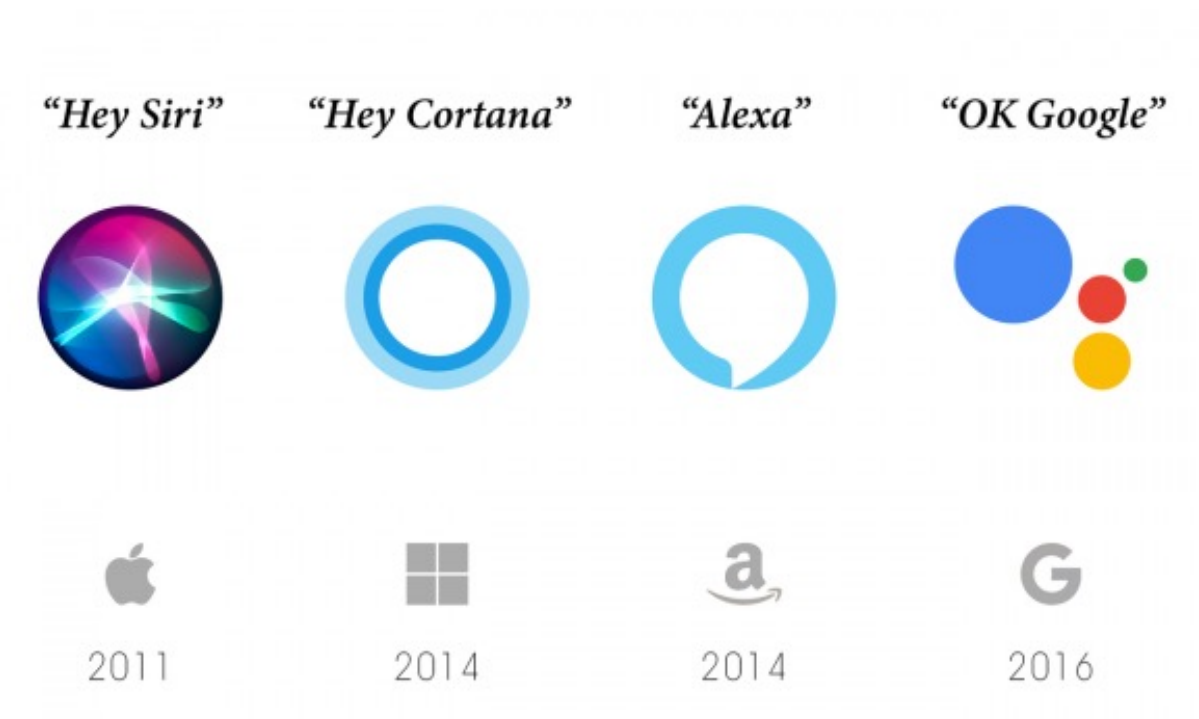
Créditos da imagem: ibm.com

# VISÃO GERAL: SINTETIZAÇÃO DA VOZ



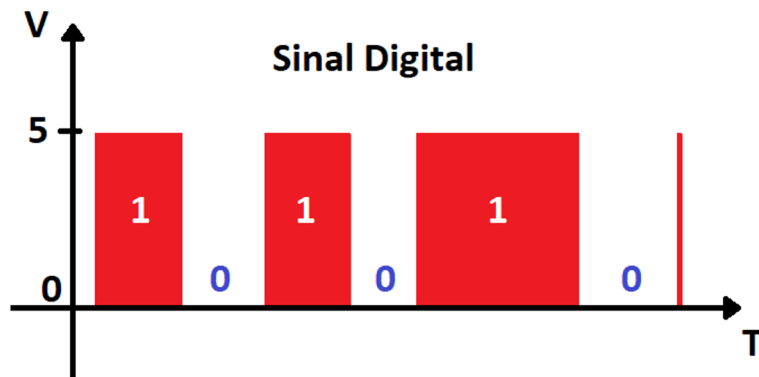
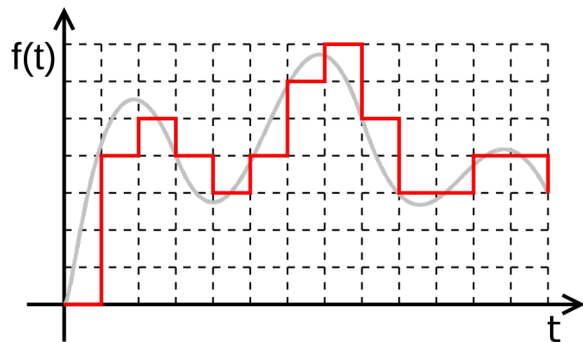
Créditos da imagem: ibm.com

# ASSISTENTES VIRTUAIS MAIS CONHECIDOS

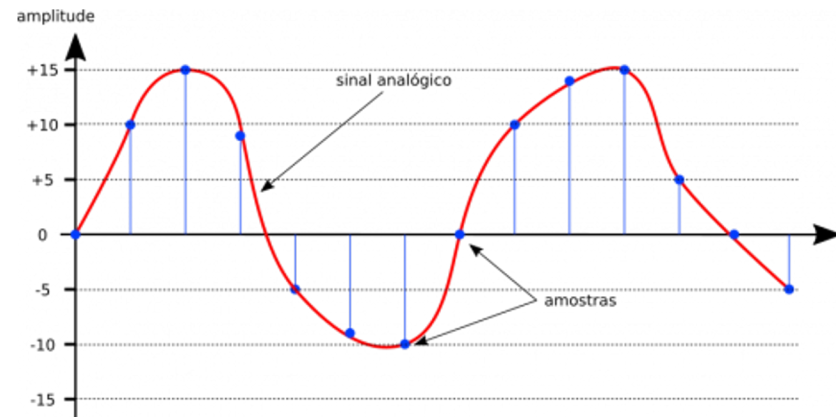


Créditos da imagem: Dribbble.com

# AMOSTRAGEM DE ÁUDIO



Créditos da imagem: embarcados.com.br



Na imagem acima, podem ser identificados o sinal analógico, proveniente de um microfone ou outro dispositivo de captura, e as amostras, que correspondem ao valor do sinal medido em intervalos regulares. O sinal analógico acima seria então representado de forma aproximada pelo seguinte vetor de amostras:

$amostras = [0, 10, 15, 9, -5, -9, -10, 0, 10, 14, 15, 5, 0, -5, \dots]$

A **taxa de amostragem** ou **frequência de amostragem** é definida como o número de amostras do sinal analógico selecionadas por segundo

# TAXA DE AMOSTRAGEM (SAMPLE RATE)

- **8000Hz**

- O padrão internacional para áudio usado em telefonia é uma taxa de amostragem de 8000 Hz (8 kHz). Isso é suficiente para que a fala humana seja compreensível

- **44100Hz – Teorema de Nyquist-Shannon**

- A taxa de amostragem de 44,1 kHz é usada para áudio de disco compacto (CD). Os CDs fornecem som estéreo de 16 bits não compactado a 44,1 kHz. O áudio do computador também usa frequentemente essa frequência por padrão

- **48000Hz**

- O áudio no DVD é gravado em 48 kHz. Isso também é frequentemente usado para áudio de computador

- **96000Hz**

- Áudio de alta resolução

- **192000Hz**

- Áudio de ultra-alta resolução. Ainda não é comumente usado, mas isso mudará com o tempo (2022)

Fonte: [developer.mozilla.org](https://developer.mozilla.org)