**UNIVERSIDADE PAULISTA**

**Murilo de Souza Ribeiro**

**Robert Danilo C. Caetano**

**Captação e análise de dados vocais, ferramenta "Áspide Recognizer”**

**RIBEIRÃO PRETO – SP**

**2019**

**Murilo Souza Ribeiro**

**Robert Danilo C. Caetano**

**Captação e análise de dados vocais, ferramenta "Áspide Recognizer”**

subtítulo

Trabalho de conclusão de curso para obtenção do título de graduação em Ciência da Computação apresentado à Universidade Paulista – UNIP

Orientador: Prof. Jean Miller Scatena

**1.INTRODUÇÃO**

Se existe algo mais fundamental do que a escrita não sabemos, mas o que realmente temos ciência é que desde os primórdios esse é o meio indispensável de registro de informações, comunicação, experiências, tratados e etc. A escrita nos agrega valores inafiançáveis, seja por meio da história e paradigmas de pessoas intelectuais que já partiram do plano terrestre, mas que deixaram legados que vêm auxiliando até o tempo contemporâneo.

Sabendo disso a computação vem nos acompanhando paralelamente cobrindo as necessidades da sociedade cada vez mais carentes, desde seu processo de evolução natural, consequentemente a décadas atrás era impensável a possibilidade de uma máquina ter a capacidade de reconhecer a voz humana e ser intérprete da sua linguagem, seja ela qual for, de modo a propiciar soluções de problemas e até mesmo mais conforto e praticidade no cotidiano. Hoje o gelo foi quebrado e a grande maioria já está adaptado com esse processo tecnológico avançado, que acabou deixando de ser novidade.

Ferramentas e recursos a todo momento são desenvolvidas e disparadas para descomplicar a vida da sociedade, focado nesse paradigma o processo de busca por um novo recurso foi pensado e conseguimos trabalhar encima das necessidades de caráter social, com um recurso primordial e fundamental, ou seja, o registro de informações de modo refinado do audível para transcrição, por esse motivo confeccionamos o software responsável pela conversão de arquivos de áudio ou captura direta através de um microfone, para formato textual.

**1.2 OBJETIVO**

**1.2.1 OBJETIVO GERAL**

O intuito deste trabalho, de forma geral, é o estudo do desenvolvimento de uma ferramenta capacitada para realizar a conversão de áudio em texto, independentemente do método utilizado na entrada de dados, seja ela por arquivo ou captura de voz.

**1.2.2 OBJETIVO ESPECÍFICO**

Dentre os objetivos destacamos:

* Estudar as formas de conversão de áudio em texto;
* Estudar as características e o funcionalidades da biblioteca Speech Recognition;
* Estudar as características e o funcionalidades do CMUSphinx;
* Desenvolver a ferramenta utilizando a linguagem Python;
* Simplificar o processo de registro de informações relevantes.

**1.3 MATERIAIS E METODOS**

Utilizaremos para a criação desta ferramenta a linguagem de programação Python, já que é uma linguagem amplamente utilizada em aplicações de reconhecimento, seja de texto, imagem ou áudio, por ser uma linguagem de fácil manipulação e escrita.

Python possui uma grande variedade de bibliotecas pra a manipulação e compreensão de áudio, nesta ferramenta serão utilizadas as bibliotecas Tkinter, o CMUSphinx.

Tkinter é uma biblioteca nativa do Python que será utilizada para a criação de interface gráfica desta ferramenta.

CMUSphinx é um conjunto d ferramentas dedicadas a compreensão da voz. Ele é composto essencialmente por quatro ferramentas: O Pocketsphinx que é a biblioteca de reconhecimento de voz; O Sphinxtrain responsável pelo treino do modelo acústico;A Sphinxbase é uma biblioteca de suporte para as demais; e por fim a Sphinx4 que é uma biblioteca de reconhecimento de voz que pode ser modificada, com seu funcionamento voltado para aplicações desktop. No CMUSphinx existem vários modelos de linguagem eainda permite a criação de novos caso seja necessário.

Nesta ferramenta utilizaremos o Pocketsphinx e a Sphinxbase para a interpretação e manipulação do áudio.