**UNIVERSIDADE PAULISTA**

**MURILO DE SOUZA RIBEIRO**

**ROBERT DANILO C. CAETANO**

**CAPTAÇÃO E ANÁLISE DE DADOS VOCAIS, FERRAMENTA "ÁSPIDE RECOGNIZER”**

**RIBEIRÃO PRETO**

**2019**

**LISTA DE ILUSTRAÇÕES**

**SUMÁRIO**

[1 INTRODUÇÃO 4](#_Toc8304979)

[**1.1 OBJETIVO 5**](#_Toc8304980)

[**1.1.1 Objetivo geral 5**](#_Toc8304981)

[**1.1.2 Objetivos específicos 5**](#_Toc8304982)

[2 ENGENHARIA DE REQUISITOS 6](#_Toc8304983)

[**2.1 Elicitação 6**](#_Toc8304984)

[**2.2 Especificação 6**](#_Toc8304985)

[**2.3 Modelagem 6**](#_Toc8304986)

[3 REFERÊNCIAS BIBLIOGRÁFICAS 7](#_Toc8304987)

# 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.

De acordo com Lima (2017) ”Uma equipe que não estabelece uma boa comunicação entre seus membros passa a impressão de estar à deriva, dando oportunidade para o erro e, consequentemente, o retrabalho”. Indubitavelmente mesmo que a tecnologia nos ajude constantemente, um dos maiores problemas dos tempos atuais, ainda é a comunicação, embora pareça simples nos comunicarmos, esse contexto vai além de uma conversa bem organizada e formal, pois a comunicação precisa ser coerente e preferencialmente que proporcione um elo eficaz entre os interlocutores, através do entendimento. E tendo em vista esse problema, propusemos que a ferramenta nos ajude também na análise das comunicações e isso contribui para o nascimento de um futuro projeto, capaz de identificar e analisar possíveis erros na comunicação entre indivíduos, através de suas expressões verbais, falta de clareza e coesão nos dialetos eletrônicos ou pessoais. E nessa aplicação pode ser encontrada a solução desse empecilho, presente grande parte das vezes em nossos relacionamentos a temível incomunicação.

## 1.1 OBJETIVO

### 1.1.1 Objetivo geral

O intuito deste trabalho, de forma geral, é o estudo e 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.

### 1.1.2 Objetivos específicos

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.

# 2 ENGENHARIA DE REQUISITOS

## 2.1 Elicitação

Através de uma discussão entre os autores deste projeto chegamos ao consenso que construiríamos uma ferramenta capaz de solucionar e auxiliar no cotidiano de quem lida constantemente com o uso de informações auditivas e tem a necessidade de torná-las mais legíveis e consequentemente facilitando a sua obtenção à análise. O seu funcionamento consiste na obtenção de informações sonoras e transcrevê -las em formato de texto.

Partindo deste princípio foi necessário idealizar como seria o seu funcionamento e quais seria os recursos utilizados em seu desenvolvimento, desde a linguagem a ser escolhida e as API's a serem utilizadas.

Após uma pesquisa de linguagens que nos propiciaria um melhor tratamento de informações brutas, concluímos que a linguagem de programação Python seria uma boa escolha, dentre as outras presentes no meio de desenvolvimento de software.

Com estudo mais aprofundados sobre os módulos da linguagem Python encontramos uma API que se encaixava perfeitamente nos parâmetros necessários para o desenvolvimento da ferramenta, chamada "SpeechRecognition". Ela nos proporciona uma variedade de recursos para transcrição de áudio em texto.

## 2.2 Especificação

## 2.3 Modelagem

## 

Capta\_audio

Def entrada\_mic

Def entrada\_file

# 3 REFERÊNCIAS BIBLIOGRÁFICAS

Lima, H. (2017). ***A Importância da Comunicação nas Organizações*.** Acesso em 31 de 03 de 2019, disponível em Endomarketing.tv: https://endomarketing.tv/importancia-da-comunicacao-nas-organizacoes/#.XKD315hKiM9

Rato, J. C. (Setembro de 2016). ***Conversação Homem-máquina****.* Acesso em 01 de 04 de 2019, disponível em IC Online: https://iconline.ipleiria.pt/bitstream/10400.8/2375/1/jo%C3%A3o%20Rato-Mestrado%20em%20Eng.Inform%C3%A1tica-Computa%C3%A7%C3%A3o%20M%C3%B3vel.pdf