transcript audio texto este projeto consiste em um programa de conversão de áudio para texto utilizando a biblioteca Speech recognition Python o objetivo principal é fornecer uma ferramenta simples e eficiente para auxiliar na transcrição de arquivos de áudio intexto a Speech recognition é uma biblioteca de Python que utiliza uma API de reconhecimento de fala para converter áudio em texto O que é uma uma API é uma interface de programação de aplicativos é um conjunto de regras e protocolos que permite que diferentes softwares se comuniquem entre si é como uma ponte que permite que o programa utilize as funcionalidades de outro de maneira padronizada e organizada essa up chamada de Google Web speech up é um serviço oferecido pela Google que faz o processamento da fala em tempo real e retorna O texto correspondente a escolha dessa biblioteca se dá pela facilidade de uso é fácil de aprender e usar o que torna a implementação do reconhecimento de fala em programa Python bastante acessível mesmo para desenvolvedores iniciantes e também por ter suporte a várias apis populares de reconhecimento de fala incluindo Google Web Speech api com recursos como as recognition desenvolvedores podem incorporar funcionalidades de processamento de linguagem natural em suas aplicações abrindo portas para interações mais intuitivas e acessíveis por meio da fala como o processo ocorre o funcionamento da Speed combine com a API de reconhecimento de fala ocorre em etapas primeiro o programa utiliza a biblioteca para ler um arquivo de áudio ou capturar áudio em tempo real em seguida a biblioteca envia esse áudio para API de reconhecimento de fala da aapi processa o áudio utilizando algoritmos avançados de reconhecimento de fala e retorna O texto correspondente a biblioteca Spirit recognition lida com toda a parte de comunicação com a api tornando o processo transparente para o desenvolvedor ela fornece métodos simples para gravar áudio enviar para a ap e receber o texto reconhecido como resposta dessa forma os desenvolvedores podem facilmente integrar o reconhecimento de fala em seus próprios programas sem se preocupar com os detalhes complexos fotos da biblioteca a biblioteca facilita a comunicação com a api permitindo que utilizemos o reconhecimento de fala em Nossos programas de maneira simples e eficiente formato de áudio de entrada neste projeto optamos pelo formato de arquivo de áudio ponto w a vi isto é aceitar a apenas entradas de arquivo w em comparação com o formato MP3 em uma ave temos uma compressão sem perdas a qualidade do som é superior ao MP3 sendo que o formato de arquivo dado luav contém muitos detalhes de som é maior que requer mais espaço porém como desejamos a mais alta qualidade e não nos importamos com o tamanho do arquivo escolhemos o formato w ave muito utilizado por profissionais de áudio bibliotecas utilizadas fizemos uns das seguintes bibliotecas nesse projeto pai sampler Speech recognition Python doc X Time e date-time vamos comentar resumidamente a função de cada biblioteca que foi importada no projeto A primeira é uma biblioteca de interface gráfica que facilita a criação de interfaces de usuário ela fornece uma API simples e intuitiva para criar janelas portões entrada e outros elementos de interface a segunda é uma biblioteca que fornece recursos de reconhecimento de fala permite captar capturar áudio ou arquivo de áudio e convertê-los em texto a terceira é uma biblioteca para manipulação de arquivo de áudio ela oferece funcionalidades para carregar cortar mesclar converter e reproduzir arquivo de áudio transformados a quarta é uma biblioteca para criação e manipulação de documentos do Microsoft Word ela permite criar editar e salvar documentos.doc além de fornecer recursos para adicionar texto formatação tabelas imagens e outros elementos aos documentos a quinta é um módulo que fornece funções relacionadas ao como medir a duração de um evento adicionar atrasos ou pausas no programa e trabalhar com time stamp A última é um módulo que fornece classe e funções para manipulação de datas e horas permite obter informações sobre a data e hora atual realizar cálculos de tempo formatar datas e horas entre outras operações relacionadas ao tempo essas bibliotecas e modos fornecem funcionalidades essenciais para o seu programa como a criação da interface gráfica reconhecimento de fala manipulação de áudio e criação de documentos do Word descrição do projeto Inicialmente pensou-se um programa que fosse simples e com poucas linhas de código assim a primeira versão do programa acabou sendo desenvolvida para lidar com a versão de pé com a conversão de pequenos trechos de áudio Ela utiliza uma função da biblioteca Recognition para realizar a conversão de áudio em texto essa versão essa versão é adequada para transcrever áudios curtos como mensagem de voz anotações curtas ou pequenos trechos de diálogos embora seja simples essa versão já pode ser útil para economizar tempo e esforço na trabalhosa transcrição manual no entanto trabalhamos com realização de entrevistas com meia hora ou mais de gravação percebemos a necessidade de uma versão mais robusto para lidar com a transcrição de áudios longos a segunda versão do programa foi tão desenvolvida para atender a essa necessidade Nessa versão o áudio dividida em pedaços de aproximadamente 30 segundos para evitar problemas de memória durante o processamento além foi realizado uma melhoria na precisão do reconhecimento de fala por meio do ajuste para o ruído ambiente antes de cada gravação daí a importância de se realizar boas gravações tentando se possível Minimizar as interferências do som ambiente pois a qualidade da gravação poderá influenciar no resultado final da impressão do texto essa versão é capaz de transcrever transcrever áudios longos como palestras e entrevistas podcast ou reuniões a relevância desse tipo de programa é notável para pessoas que trabalham com a conversão manual de arquivos para texto transcrever áudios manualmente pode ser um processo demorado e pede oso exigindo muita atenção e esforço ouvir cada frase e em seguida escrevê-las com auxílio de um programa de conversão automática como este o trabalho é facilitado e acelerado permitindo que as pessoas fogem em outras tarefas importantes em resumo esse programa de conversão de áudio em texto é uma ferramenta valiosa para economizar tempo e esforço na transcrição manual as duas versões voltar para textos curtos e a segunda para textos longos fornecem soluções específicas para diferentes necessidades através desse projeto buscamos fornecer uma solução eficiente e acessível para melhorar a produtividade e facilitar a vida daqueles que precisam lidar com transcrição de