



# Whatsapp Transcriber

Acelere as investigações com reconhecimento de fala

Desenvolvido por Cristiano Ritta, Delegado de Polícia do Estado do Rio Grande do Sul



## Reconhecimento de fala e conversão em texto

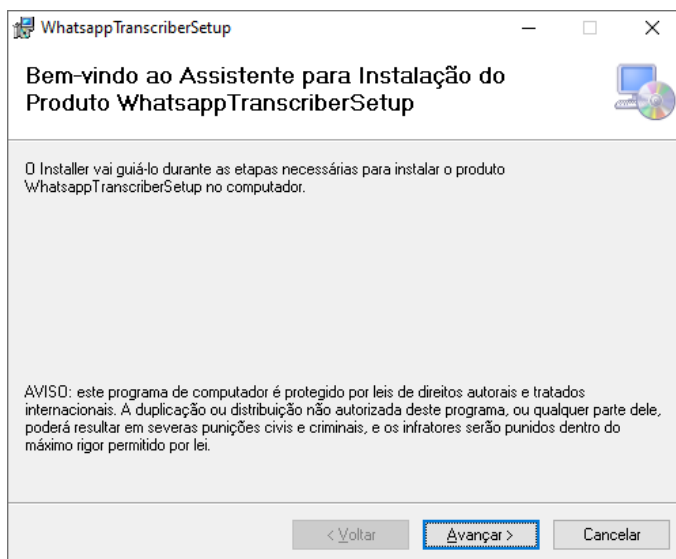
O reconhecimento de fala é uma área da computação que estuda a linguística para desenvolver métodos e tecnologias que permitam o reconhecimento e a transcrição de linguagem falada de maneira automática.

O Whatsapp Transcriber foi desenvolvido usando um poderoso motor de transcrição para auxiliar os profissionais que precisam processar e analisar milhares de arquivos de áudio do comunicador Whatsapp. O software reconhece o formato do arquivo, implementa o codec necessário para que ele possa ser ouvido em qualquer computador, processa o conteúdo e converte em texto em português, além de permitir que sejam feitas pesquisas nos áudios para encontrar expressões ou palavras relevantes para a investigação.

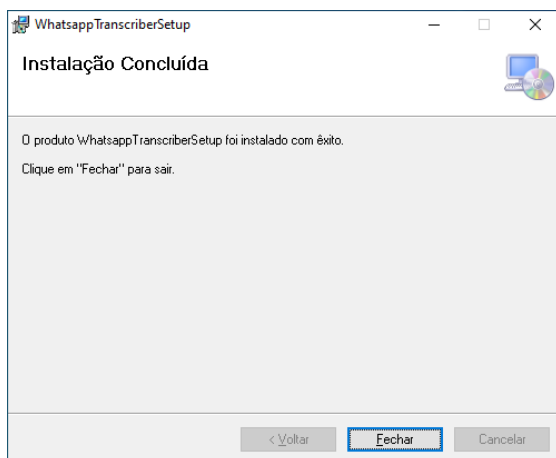
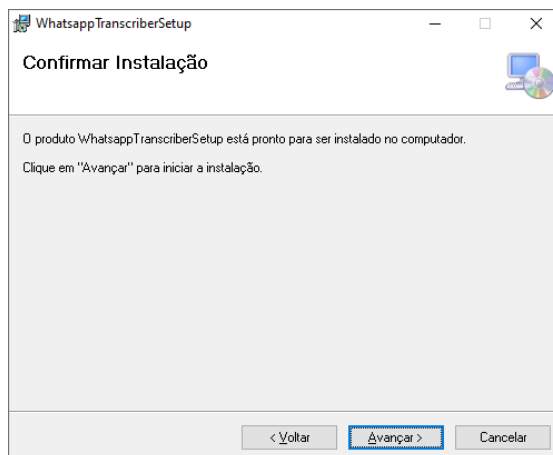
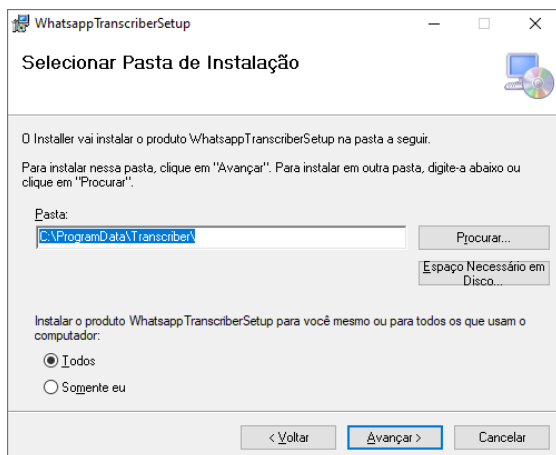
# 1. Instalação

Para instalar o software, faça o download em <https://github.com/delcristianoritta/whatsapptranscriber>

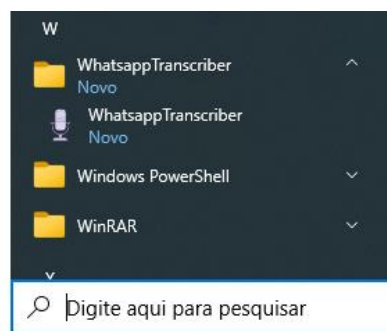
O Assistente de instalação guiará você por todo o processo.



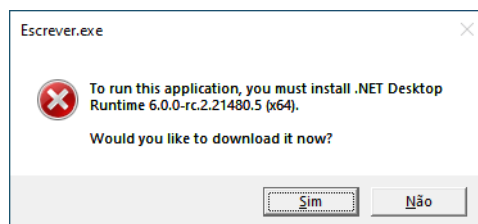
Instale o sistema na pasta recomendada **C:\ProgramData\Transcriber\**



Ao concluir a instalação serão criados os atalhos na Área de Trabalho e no Menu Iniciar > Whatsapp Transcriber



Ao executar o arquivo o sistema vai procurar por uma biblioteca de dados do Windows (.Net v.6.0.6). Essa biblioteca é necessária para o correto funcionamento do software. Se não estiver instalada em seu computador o sistema vai orientá-lo em como fazer o download ([Link para download](#)).



## 2. Executando o Software



# 2.1 Passo a Passo

1

## SELECIONAR PASTA

Se você deseja transcrever todos os áudios de uma só vez, clique no botão “Procurar” e escolha a pasta que contém os arquivos de áudio .opus (geralmente esses arquivos ficam na pasta Whatsapp\Media\Whatsapp Voice Notes). Não é preciso selecionar as subpastas individualmente, caso você deseje processar todas elas de uma só vez. O sistema reconhecerá as subpastas e processará cada uma delas.

2

## CONTADOR DE ARQUIVOS

Quando você selecionar a pasta o programa vai reconhecer quantos arquivos de áudio foram encontrados. Ao longo do processo de transcrição automática o contador de “arquivos transcritos” será atualizado de acordo com o progresso.

3

## INICIAR

Clique no botão “Iniciar” para começar o processo. Todas as tarefas serão realizadas automaticamente. Cada arquivo de texto (.txt) da transcrição será salvo junto com o arquivo de áudio original. Além disso, o sistema vai gerar um arquivo em formato HTML (relatório.html) na pasta selecionado, com o conteúdo de todas as transcrições, facilitando a busca por palavras chaves e falas específicas nos diálogos.

Durante o progresso o contador de arquivos (2) e a barra de progresso (4) serão atualizados para facilitar o acompanhamento da tarefa. Você pode parar a transcrição clicando no botão “Parar”.

5

## TRANSCREVER ARQUIVO

Se você desejar transcrever um único arquivo, clique no botão “Transcrever arquivo” e selecione o áudio. O programa vai salvar o resultado em um arquivo de texto com o mesmo nome e pasta do original.

Arquivos	Transcrição
Arquivo: PTT-20211012-WA0000.opus Caminho: \20211012\PTT-20211012-WA0000.txt Data: 07/02/2022 16:02:39 Tamanho (bytes): 157 MD5: 0f0963ee8a324a431012567f5dc8a9308	Olá vindo aqui fala aqui dá para ir de carro não vai ter mais nada aí tem que que eu tenho que chegar aqui uma coisa que é trem-bala também quer uma balaína
Arquivo: PTT-20211012-WA0005.opus Caminho: \20211012\PTT-20211012-WA0005.txt Data: 07/02/2022 16:02:39 Tamanho (bytes): 66 MD5: 837e4970228262183030F34617C560e	se não for assim nem venho por que negócio negócio negócio
Arquivo: PTT-20211012-WA0010.opus Caminho: \20211012\PTT-20211012-WA0010.txt Data: 07/02/2022 16:02:39 Tamanho (bytes): 43 MD5: ca18a2d99905381e3540eb8d2878c15	não que eu vá hoje mas onde é que é um nó
Arquivo: PTT-20211012-WA0013.opus Caminho: \20211012\PTT-20211012-WA0013.txt Data: 07/02/2022 16:02:39 Tamanho (bytes): 34 MD5: 26a891191d88237899ab937c7c7c	eu tava dormindo acabar de acordar
Arquivo: PTT-20211012-WA0015.opus Caminho: \20211012\PTT-20211012-WA0015.txt Data: 07/02/2022 16:02:39 Tamanho (bytes): 65 MD5: 7c05c0f0209342430f031968ba1854	é não que eu vá também né mas eu também queria saber onde é que é
Arquivo: PTT-20211012-WA0016.opus Caminho: \20211012\PTT-20211012-WA0016.txt Data: 07/02/2022 16:02:39 Tamanho (bytes): 28 MD5: 98582d569e0e0e0e3d5c70675201798b	arrasta para cima toca Negra
Arquivo: PTT-20211012-WA0017.opus Caminho: \20211012\PTT-20211012-WA0017.txt Data: 07/02/2022 16:02:39 Tamanho (bytes): 134 MD5: 8881c0fed30b0a099d76d22616d31	Quem disse que me superou Deus a volta por cima e quer saber da minha vida arrasta para cima quem não tem mais amor
Arquivo: PTT-20211012-WA0018.opus Caminho: \20211012\PTT-20211012-WA0018.txt Data: 07/02/2022 16:02:39 Tamanho (bytes): 232 MD5: 1a22a08b8a6a26028845a7070f2f68	vou ficar mais um tempo aqui que agütar a mãe a mãe tá com problema aqui com mais com tu eu vou ficar por aqui vou agütar ela depois sim aí quando tiver tudo tranquilo aqui aí eu vou pa vem domingo também pode vir para lá para casa

(exemplo de relatório da transcrição)

## Copyright

O Whatsapp Transcriber é uma ferramenta gratuita, desenvolvida para auxiliar policiais, peritos e membros do Ministério Público com a tarefa de transcrever e analisar milhares de arquivos de mensagem de voz no mensageiro Whatsapp.

Se você deseja ajudar a aprimorar o software, entre em contato:

Desenvolvido por

**Cristiano Ribeiro Ritta**

Delegado de Polícia do Estado do Rio Grande do Sul  
cristiano-ritta@pc.rs.gov.br

<https://br.linkedin.com/in/cristiano-ribeiro-ritta-7416ab27>



# Whatsapp Transcriber

Acelere as investigações com reconhecimento de fala