

OpenAI Whisper

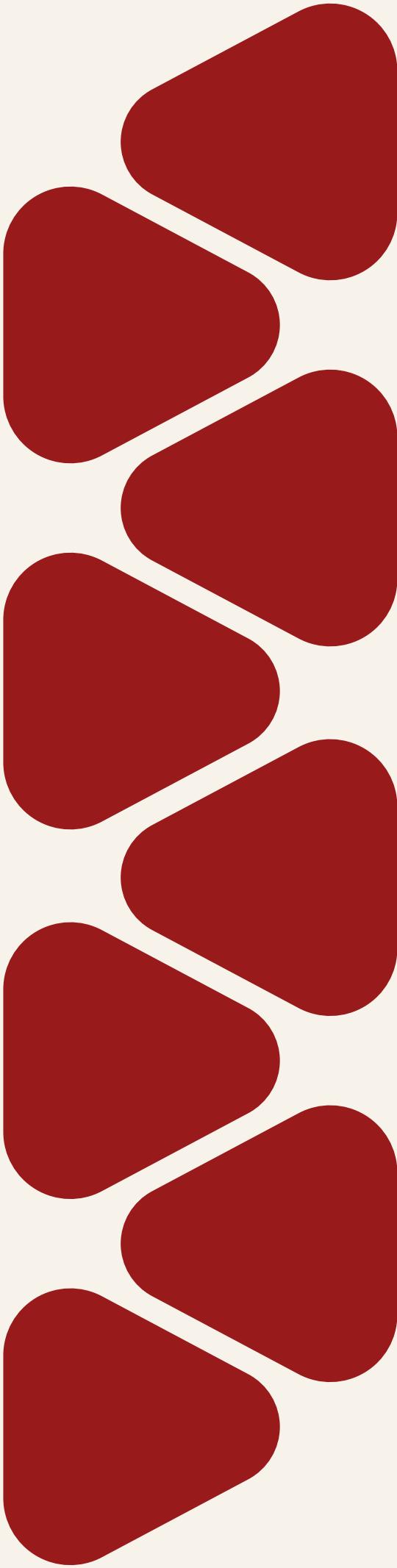
**transcrição
automática
para pesquisa**

**Guia Prático de
Utilização do
OpenAI Whisper no
Google Colab**

Thales Morbach Lange



Grupo de Pesquisa
Ideologia e Análise de Discurso
PPGCPol - UFPel



OpenAI Whisper

transcrição automática
para pesquisa

Guia Prático de Utilização do
OpenAI Whisper no Google Colab

Thales Morbach Lange



Grupo de Pesquisa
Ideologia e Análise de Discurso
PPGCPol - UFPel

Capa:

Corina Gonsales Vega

Laura da Silva Rhoden Machado

Thales Morbach Lange

Diagramação:

Thales Morbach Lange

Dados de catalogação da Publicação
Fabiano Domingues Malheiro – CRB 10/1955

L274o Lange, Thales Morbach

OpenAI Whisper [recurso eletrônico]: transcrição automática para pesquisa (Guia Prático de Utilização do OpenAI Whisper no Google Colab) / Thales Morbach Lange. — Pelotas: PPGCPol, 2025.

25 p. : il. : PDF; 4,43 MB

ISBN: 978-65-01-53228-8

1. Google Colab.
2. Estrutura de código.
3. Modelo Whisper.
4. OpenAI Whisper I. Título.

CDD:005.133

SUMÁRIO

1. Introdução	4
2. Pré-requisitos.....	5
3. Acessando o projeto	6
3.1. Site do GP IdAD	6
3.2. Google Colab.....	7
4. Estrutura do projeto e funcionamento do Colab.....	8
4.1. Lógica de funcionamento e estrutura do código.....	9
5. Executando o	10
5.1. Célula 1, dependências e bibliotecas.....	10
5.2. Célula 2, modelo Whisper.....	11
5.3. Célula 3, idioma.....	12
5.4. Célula 4, transcrição	14
5.5. Célula 5, transcrição em lotes.....	16
5.6. Célula 6, ferramenta extra: yt-dlp	17
6. Dicas úteis.....	19
6.1. Salvando os arquivos gerados	19
6.2. Limpeza e organização da interface	19
6.3. Reutilizando o projeto	20
6.4. Recuperando o notebook original.....	20
7. Erros comuns e soluções	21
8. (Algumas) Considerações finais.....	22
9. Notas do desenvolvedor.....	23
9.1. Nota sobre o desenvolvimento dos códigos.....	23
9.2. Nota ao usuário: adaptação livre e criativa do código	23
10. Apêndice: Checklist para auxílio no passo a passo da execução do Whisper.....	25

1. INTRODUÇÃO

A análise de pronunciamentos, entrevistas e outros registros em áudio e/ou vídeo é fundamental para pesquisas em Ciência Política e áreas afins das humanidades. Tradicionalmente, a transcrição manual desses materiais demanda tempo e recursos, limitando a escala das análises. O uso de ferramentas de Inteligência Artificial (IA), como o Whisper, aliado à acessibilidade do Google Colab, oferece uma solução metodológica inovadora, permitindo a transcrição automática de grandes volumes de dados qualitativos com alta precisão. Essa ferramenta não apenas otimiza processos, mas também amplia as possibilidades de investigação crítica, garantindo reproduzibilidade e rigor metodológico.

O **OpenAI Whisper¹** é um modelo de transcrição automática de áudio de alta precisão, suportando múltiplos idiomas. Este guia prático ensina a utilizá-lo no **Google Colab**, uma plataforma gratuita que permite executar código Python sem instalação local.

O Google Colab (abreviação de Colaboratory) é uma plataforma gratuita oferecida pelo Google que permite executar códigos em Python diretamente no navegador, sem precisar instalar nada no seu computador. Ele foi criado especialmente para atividades educacionais, científicas e de prototipagem de projetos com inteligência artificial, como o Whisper.

Uma das grandes vantagens do Colab é que ele já vem com recursos prontos para uso, como acesso a placas de vídeo (GPU) e integração com o Google Drive, facilitando a manipulação de arquivos. Além disso, permite compartilhar projetos facilmente com outras pessoas.

¹ Confira a documentação do OpenAI Whisper em: <https://github.com/openai/Whisper>.

2. PRÉ-REQUISITOS

Antes de iniciar a execução do projeto, é importante garantir que você tenha algumas condições mínimas para que tudo funcione corretamente. Esses pré-requisitos não exigem conhecimentos técnicos avançados, apenas recursos básicos que ajudam a evitar erros durante a execução. Ter uma conta Google ativa, uma conexão estável com a internet e os arquivos nos formatos certos já são suficientes para a maioria das pessoas conseguirem utilizar a ferramenta com sucesso. A seguir, listamos o que é necessário — e o que pode facilitar ainda mais sua experiência.

- Conta Google (para acessar o Colab e Drive).
- Recomenda-se usar Google Chrome ou Firefox atualizado para evitar erros de compatibilidade.
- Conexão estável com a internet.
- Arquivos de áudio ou vídeo em formatos compatíveis (.mp3, .wav, .m4a, etc.).
- Conhecimento básico de navegação em pastas (opcional, mas útil).
- Saber como renomear e localizar um arquivo no computador pode ajudar.

Formatos de arquivo aceitos pelo Whisper:

.mp3, .mp4, .m4a, .wav, .webm, .flac, .ogg, .aac, .wma, .mkv, entre outros formatos populares de áudio e vídeo.

3. ACESSANDO O PROJETO

Antes de iniciar o uso da ferramenta, é necessário abrir o arquivo do projeto no ambiente adequado. Essa etapa consiste basicamente em dois momentos: primeiro, fazer o download do arquivo de código já pronto (um tipo especial de arquivo chamado *notebook*), e depois carregá-lo na plataforma **Google Colab**, que é onde o código será executado. O processo é simples e não exige nenhum programa instalado no computador, apenas acesso à internet e uma conta Google.

Mas antes de começar, é importante entender brevemente o que é esse tipo de arquivo. O Whisper será executado dentro de um **notebook Jupyter**, um arquivo com extensão `.ipynb`. Esse notebook funciona como um “caderno interativo”, que combina explicações em texto com blocos de código que podem ser executados individualmente. Isso torna o processo mais organizado e acessível, especialmente para quem está aprendendo ou deseja seguir um passo a passo bem estruturado.

No caso deste projeto, utilizaremos o notebook *Whisper_IdAD_5.0².ipynb*, desenvolvido e disponibilizado previamente com todas as instruções, códigos e configurações necessárias para executar o Whisper no Google Colab. A seguir, explicamos como acessar e abrir esse arquivo para começar a transcrição automática de áudios e vídeos.

3.1. Site do GP IdAD

O primeiro passo é acessar o arquivo "Whisper_IdAD_5.0.ipynb". Para isso:

→ Acesse o site do GP IdAD:

<https://wp.ufpel.edu.br/idad/produtos/produtos-tecnicos/>

² A extensão "_5.0" (ou outro número) na nomenclatura do arquivo é apenas para critério de controle do desenvolvedor, visto que a estrutura do código passou por modificações desde a sua criação e continua suscetível a novas modificações, caso necessário.

Vá até o link do arquivo, clique nele e o download será iniciado automaticamente.

Após realizar o download do arquivo a partir do site do GP IdAD, ele virá compactado em formato `.zip`. Esse tipo de arquivo serve para agrupar e reduzir o tamanho de outros arquivos. Para poder utilizar o projeto, é necessário **extraír** o conteúdo desse `.zip`. Para isso, basta localizar o arquivo na sua pasta de downloads, clicar com o botão direito sobre ele e escolher a opção “**Extraír tudo...**” (em computadores com Windows) ou “**Descompactar**” (em sistemas como macOS ou Linux). Após a extração, será criada uma nova pasta com o arquivo `Whisper_IdAD_5.0.ipynb` dentro dela. Esse é o arquivo que você irá carregar no Google Colab na próxima etapa.

3.2. Google Colab

Para executar o projeto:

→ Acesse o Google Colab: <https://colab.research.google.com>

→ Clique em *Upload > Procurar* e selecione o arquivo `.ipynb` que você extraiu.

O projeto será carregado e armazenado no seu Google Drive.

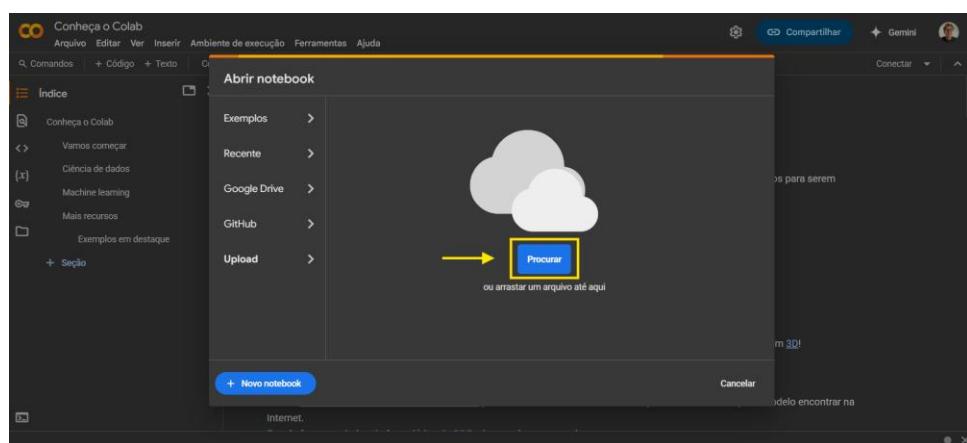


Figura 1- Tela de seleção de arquivo no Google Colab

4. ESTRUTURA DO PROJETO E FUNCIONAMENTO DO COLAB

Antes de começar, é necessário se familiarizar com a interface do projeto e o funcionamento do Google Colab para saber como executar a ferramenta corretamente.

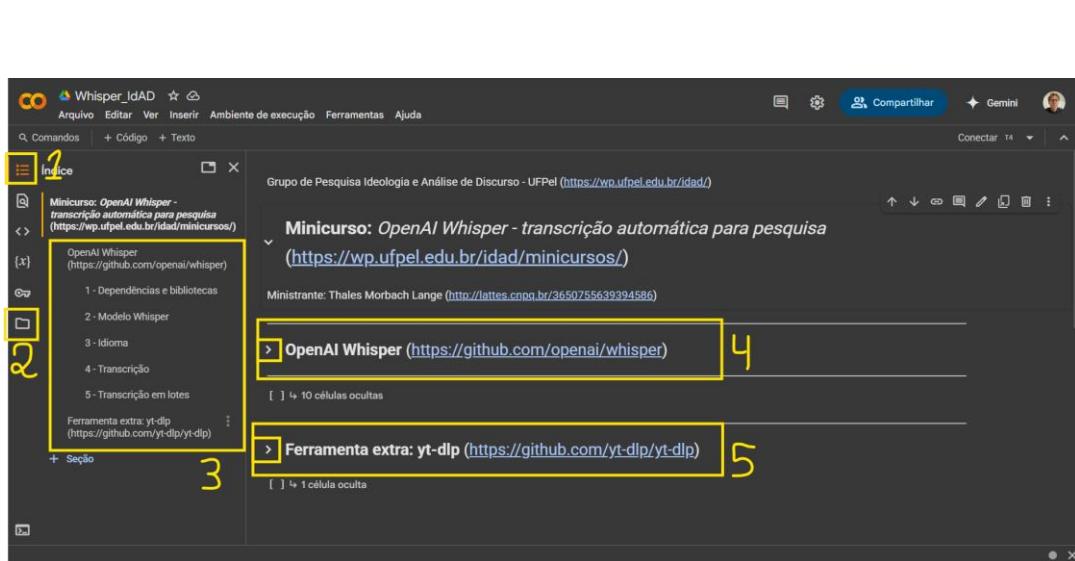


Figura 2- Apresentação da interface e recursos do Google Colab

Ao abrir o projeto no Google Colab, você verá uma interface organizada em seções (itens 4 e 5 da Figura 2) que facilitam a navegação e a execução do código. No canto superior esquerdo, há um menu lateral com dois ícones importantes: o primeiro (ícone de lista, item 1 da Figura 2) exibe o **sumário do notebook** (Item 3 da Figura 2), permitindo que você pule diretamente para qualquer parte do código com um clique. O segundo ícone (pasta, item 2 da Figura 2) abre o **painel de arquivos**, onde você pode enviar, visualizar, renomear ou excluir arquivos, além de acessar os resultados gerados pelas transcrições. Ao centro da tela ficam as **células de código**, cada uma com um botão de "play" (▶) no canto esquerdo — esse botão serve para **executar o código contido naquela célula**. A interface foi pensada para ser simples e funcional, permitindo que mesmo quem não tem experiência em programação consiga acompanhar o passo a passo com tranquilidade. A imagem nesta seção do material ilustra onde cada um desses elementos está localizado na tela.

Outro ponto importante é que o código do projeto já está configurado para montar automaticamente o Google Drive, ou seja, você não precisa digitar nenhum comando para isso — ao executar a primeira célula, o Colab irá solicitar uma autorização. Ao permitir o acesso, seu Drive será vinculado ao ambiente, e você poderá salvar ou acessar arquivos diretamente da sua conta Google, facilitando muito o processo. Com isso, toda a transcrição pode ser feita e armazenada na nuvem, sem depender da memória local do seu computador.

9

4.1. Lógica de funcionamento e estrutura do código

O funcionamento do projeto Whisper no Google Colab se organiza em etapas chamadas **células**, que são como pequenos blocos de instruções executadas separadamente. Cada célula tem uma função específica dentro do processo de transcrição. A **Célula 1** é responsável por instalar e carregar todas as bibliotecas e ferramentas que o Whisper precisa para funcionar. Já a **Célula 2** permite que o usuário escolha qual modelo do Whisper será utilizado — desde versões mais rápidas e leves até as mais precisas e completas. A seguir, a **Célula 3** define o idioma do conteúdo a ser transcreto, o que é essencial para garantir um bom reconhecimento da fala.

A **Célula 4** realiza a transcrição de **um único arquivo por vez**, sendo ideal para quem vai trabalhar com áudios ou vídeos isolados. Por outro lado, a **Célula 5** permite a **transcrição em lote**, ou seja, de vários arquivos de uma só vez — uma funcionalidade útil para quem tem muitos materiais a transcrever e quer automatizar esse processo. Importante: **as células 4 e 5 são independentes entre si** — você pode usar apenas uma delas, conforme sua necessidade.

Por fim, a **Célula 6** é uma ferramenta opcional e complementar: ela usa o programa yt-dlp para **baixar vídeos da internet diretamente para o ambiente do Colab**, como por exemplo do YouTube. Assim como as anteriores, **ela também funciona de forma independente** e pode ser utilizada antes de qualquer transcrição, caso você deseje transcrever um conteúdo que ainda não está salvo no seu computador.

5. EXECUTANDO O WHISPER

O Google Colab funciona como um computador remoto temporário. Sempre que abrimos o projeto, estamos acessando uma nova máquina virtual, o que significa que todos os arquivos e bibliotecas precisam ser instalados novamente a cada uso.

10

```
# Célula 1: Instalação das dependências e importação das bibliotecas necessárias
!pip install git+https://github.com/openai/whisper.git
!pip install torch torchaudio
!sudo apt update && sudo apt install ffmpeg
!pip install ipywidgets
!pip install google-colab

# Importações de bibliotecas
import os
import whisper
import ipywidgets as widgets
import numpy as np
from IPython.display import display, clear_output, Audio
from google.colab import drive, files, auth
from googleapiclient.discovery import build
from googleapiclient.http import MediaFileUpload
import datetime

print("✅ Bibliotecas instaladas e importadas com sucesso!")
```

Figura 3- Demonstração de funcionamento das 'células' no Google Colab

Como mencionado antes, seu funcionamento se dá através das células de código, cada uma com sua função, e executadas em uma determinada ordem. Cada célula possui um botão de play (►) em seu canto superior esquerdo, que é responsável por executar o código contido nela.

5.1. Célula 1, dependências e bibliotecas

A partir disso, accese a célula 1 e clique no botão de play, que fica no canto superior esquerdo de cada célula. O procedimento de instalação dos arquivos necessários leva em média 2 minutos para ser concluído. Ao ser concluído esse processo, aparece a seguinte mensagem na saída da célula

(parte inferior da célula onde aparecem os resultados da execução do código):

" Bibliotecas instaladas e importadas com sucesso!"

Importante:

→ Permissões do Google Drive: Durante a execução desta célula, irá aparecer uma solicitação: "Permitir que este notebook acesse seus arquivos do Google Drive?"

Clique em Permitir para que o Whisper possa acessar e salvar arquivos no seu Drive. Essa permissão é segura e temporária, válida apenas para a sessão atual do Colab.

Depois disso, a célula pode ser recolhida para não poluir a interface, pois ela não será mais usada na execução da ferramenta.

5.2. Célula 2, modelo Whisper

O próximo passo é selecionar o modelo Whisper. No Whisper, o **modelo** é a inteligência artificial responsável por "ouvir" o áudio e transformá-lo em texto. Ele é o núcleo do sistema de transcrição e foi treinado com milhares de horas de falas humanas em diferentes idiomas. Existem modelos de diferentes tamanhos, cada um com uma capacidade de compreensão, velocidade de processamento e nível de precisão diferentes.

Você escolhe o modelo de acordo com suas necessidades: modelos pequenos são rápidos e leves, enquanto modelos maiores são mais lentos, mas entregam transcrições mais completas e confiáveis.

O Whisper tem cinco principais modelos de funcionamento:

Modelo	Velocidade	Quando usar
"tiny"	Muito rápido	Bom para testes rápidos. Baixa precisão. Ideal em computadores fracos.
"base"	Muito rápido	Levemente mais preciso que o tiny. Útil para áudios bem limpos.

"small"	Rápido	Precisão mediana, ainda rápido. Funciona bem para áudios mais limpos.
"medium"	Moderadamente Lento	Equilíbrio entre tempo e qualidade. Ideal para pesquisas sérias.
"large"	Mais lento	Alta qualidade, entende melhor áudios difíceis. Pode demorar mais.

Para configurar o modelo Whisper é preciso navegar até a terceira linha da célula 2, e digitar o modelo a ser usado entre as aspas – sempre se atente para isso –, como ilustrado na imagem abaixo.



```
[ ] # célula 2: Seleção do modelo Whisper
      # Configuração do modelo
→ modelo = 'medium' # Pode ser: "tiny", "base", "small", "medium", "large"
      ↑
      # Carregar o modelo
      print(f"⏳ Carregando modelo {modelo}...")
      model = whisper.load_model(modelo)
      print("✅ Modelo carregado e pronto para uso!")
```

Figura 4- Demonstração de funcionamento da Célula 2

Após indicar o modelo a ser usado, aperte o botão de play da célula e aguarde a indicação de quando o modelo estiver carregado e pronto para uso. Irá aparecer a mensagem:

"✅ Modelo carregado e pronto para uso!"

5.3. Célula 3, idioma

Após configurar o modelo Whisper que operacionalizará a transcrição, precisamos fazer a indicação do idioma do material que queremos

transcrever, visto que o Whisper suporta uma grande quantidade de idiomas diferentes.

Os principais idiomas suportados pelo Whisper são:

Idioma	Sigla/Código
Auto (Detectar automaticamente) ³	None
Inglês	"en"
Português	"pt"
Espanhol	"es"
Francês	"fr"
Alemão	"de"
Italiano	"it"
Holandês	"nl"
Russo	"ru"
Japonês	"ja"
Chinês (Mandarim)	"zh"
Coreano	"ko"
Árabe	"ar"
Hindi	"hi"

Para configurar o idioma, seguiremos praticamente os mesmos passos da seleção do modelo Whisper: precisamos digitar a sigla/código do idioma desejado entre as aspas indicadas na terceira linha da célula 3, como indicado abaixo:

³ Neste caso específico, utilize None sem as aspas e com o primeiro N maiúsculo.

▼ 3 - Idioma

```
[ ] # Célula 3: Idioma

# Configuração do idioma para transcrição
→ idioma = "pt" # Pode ser: None (auto), "en", "pt", "es", "fr", "de", etc.
    ↑
# Dicionário de referência (apenas para consulta)
IDIOMAS_SUPORTADOS = {
    None: "Auto (Detectar automaticamente)",
    "en": "Inglês", "pt": "Português", "es": "Espanhol",
    "fr": "Francês", "de": "Alemão", "it": "Italiano",
    "nl": "Holandês", "ru": "Russo", "ja": "Japonês",
    "zh": "Chinês (Mandarim)", "ko": "Coreano",
    "ar": "Árabe", "hi": "Hindi"
}

# Exibir configuração
print("\n🌐 Idioma selecionado para transcrição:")
print(f"→ {IDIOMAS_SUPORTADOS.get(idioma, 'Desconhecido')}")
if idioma is None:
    print("(O idioma será detectado automaticamente)")
```

Figura 5- Demonstração de funcionamento da Célula 3

Após indicar o modelo a ser usado é só apertar o botão de play da célula e aguardar a confirmação da seleção do idioma na saída da célula.

5.4. Célula 4, transcrição

A célula 4 é responsável por realizar a transcrição propriamente dita do arquivo de áudio. Antes de executá-la, é necessário configurar alguns parâmetros:

→ Caminho do arquivo de áudio: No parâmetro "arquivo_audio", substitua "/content/audio.mp3" pelo caminho completo do arquivo que deseja transcrever. Certifique-se de que o arquivo está acessível no ambiente do Colab, seja através do upload direto ou do Google Drive.

Para obter o caminho do arquivo, acesse o painel de arquivos e navegue pelas pastas até localizar o arquivo desejado. Ao localizar o arquivo, clique com o botão direito do mouse sobre ele e selecione a opção "Copiar caminho".

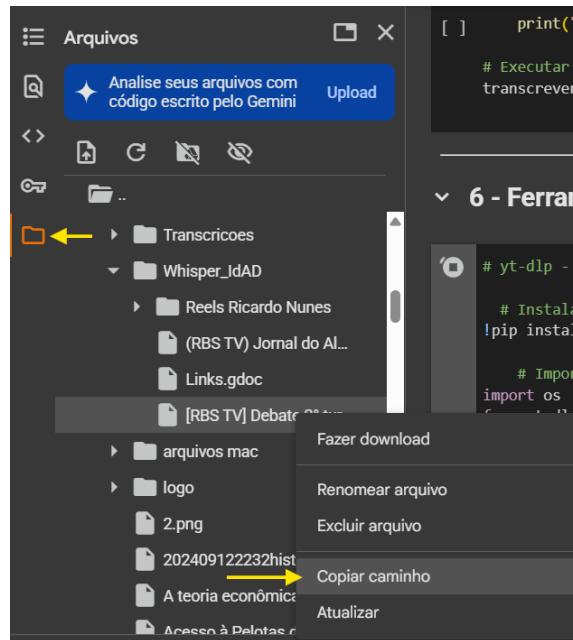


Figura 6- Demonstração de funcionamento da seção de arquivos do Google Colab

Após isso, basta colar o local do arquivo no parâmetro indicado acima, sempre tomando cuidado para ficar dentro das aspas.

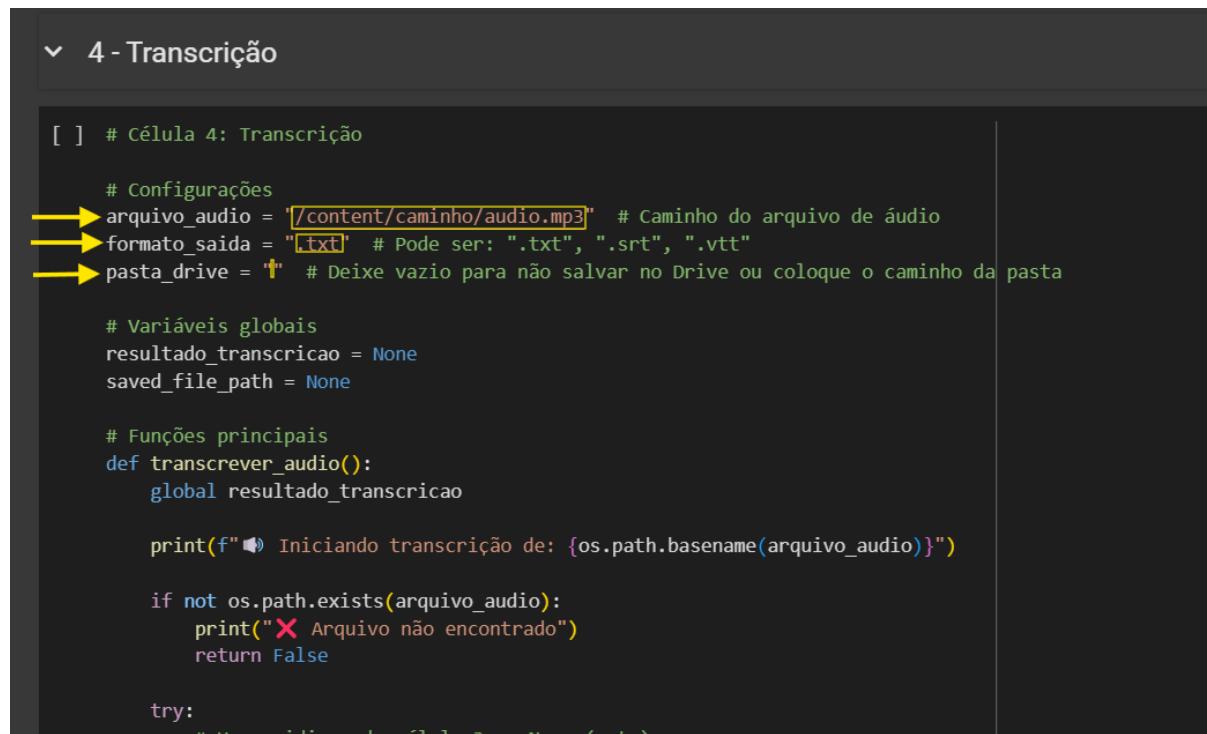
→ Formato de saída: No parâmetro "formato_saida" escolha o formato do arquivo de transcrição entre as opções:

".txt": Arquivo de texto simples.

".srt": Arquivo de legenda com timestamps (ideal para vídeos).

".vtt": Similar ao .srt, mas em formato WebVTT.

→ Pasta do Drive: Caso queira salvar a transcrição diretamente no Google Drive, no parâmetro "pasta_drive", insira o caminho da pasta de destino, ou deixe vazio para salvar apenas localmente no Colab.



```
[ ] # Célula 4: Transcrição

# Configurações
arquivo_audio = '/content/caminho/audio.mp3' # Caminho do arquivo de áudio
formato_saida = ".txt" # Pode ser: ".txt", ".srt", ".vtt"
pasta_drive = '' # Deixe vazio para não salvar no Drive ou coloque o caminho da pasta

# Variáveis globais
resultado_transcricao = None
saved_file_path = None

# Funções principais
def transcrever_audio():
    global resultado_transcricao

    print(f"▶ Iniciando transcrição de: {os.path.basename(arquivo_audio)}")

    if not os.path.exists(arquivo_audio):
        print("✗ Arquivo não encontrado")
        return False

    try:
        # Usar a edição da célula 3 ou None (auto)
        resultado_transcricao = whisper.transcribe(arquivo_audio, formato=formato_saida)
        saved_file_path = resultado_transcricao['path']
        print(f"✓ Transcrição concluída! Salvo em: {saved_file_path}")
    except Exception as e:
        print(f"✗ Erro durante a transcrição: {e}")

    return resultado_transcricao
```

Figura 7- Demonstração de funcionamento da Célula 4

Após configurar os parâmetros, clique no botão de *play* para executar a célula. O processo de transcrição pode levar alguns minutos, dependendo do tamanho do arquivo e do modelo Whisper selecionado. Ao final, você verá uma mensagem de confirmação e o caminho onde o arquivo de transcrição foi salvo.

5.5. Célula 5, transcrição em lotes

A célula 5 permite transcrever múltiplos arquivos de áudio de uma só vez, o que é ideal para projetos com grande volume de dados. Para utilizá-la:

→ Pasta de origem: Insira o caminho da pasta no Google Drive onde os arquivos de áudio estão armazenados. Por Exemplo:

`"/content/drive/MyDrive/pasta".`

→ Pasta de destino: Defina o caminho da pasta onde as transcrições serão salvas. Certifique-se de que a pasta existe ou deixe que o script a crie automaticamente.

▼ 5 - Transcrição em lotes

```

▶ # Célula 5: Transcrição em lotes com Google Drive

# Configurações
pasta_origem = '/content/drive/MyDrive/caminho/pasta'
pasta_destino = "/content/drive/MyDrive/caminho/pasta"

# Montar Google Drive
print("\n⌚ Montando Google Drive...")
from google.colab import drive
drive.mount('/content/drive')
print("✅ Google Drive montado com sucesso!")

# Função para transcrever em lote
def transcrever_lote(origem, destino):
    # Criar pasta de destino se não existir
    os.makedirs(destino, exist_ok=True)

    # Listar arquivos de áudio
    arquivos = [f for f in os.listdir(origem)
                if f.lower().endswith('.mp3', '.wav', '.m4a', '.ogg', '.flac', '.mp4', '.mpeg')]

    if not arquivos:
        print("❌ Nenhum arquivo de áudio encontrado na pasta de origem!")

```

Figura 8- Demonstração de funcionamento da Célula 5

Ao executar a célula, o script listará todos os arquivos de áudio na pasta de origem e iniciará o processo de transcrição um por um. O progresso será exibido para cada arquivo, incluindo o idioma selecionado e o status da transcrição. Os arquivos de saída serão salvos no formato `.txt` e nomeados como:

`"[nome_do_arquivo]_transcricao.txt"`.

5.6. Célula 6, ferramenta extra: yt-dlp

O projeto inclui uma ferramenta adicional chamada `yt-dlp`⁴, que permite baixar áudio de vídeos do YouTube e de outros sites da internet diretamente no Colab. Para utilizá-la:

⁴ Confira a documentação do `yt-dlp` em: <https://github.com/yt-dlp/yt-dlp>

→ Instalação: A célula já inclui o comando para instalar o *yt-dlp*. Basta executá-la uma vez.

→ Download de áudio: Cole a URL do vídeo no campo indicado e aperte *Enter*. O áudio será baixado no formato MP3 e salvo no ambiente do Colab, pronto para ser transcrita pelo Whisper.

Durante o download, o progresso será exibido, incluindo a velocidade de download e o tempo restante.

```
print("YT-DLP no Google Colab - Download para MP3")
print("-----")
url = input("Cole a URL do vídeo que deseja baixar como MP3: ")
baixar_mp3_com_progresso(url)

...
Collecting yt-dlp
  Downloading yt_dlp-2025.5.22-py3-none-any.whl.metadata (174 kB)
  174.3/174.3 kB 10.0 MB/s eta 0:00:00
  Downloading yt_dlp-2025.5.22-py3-none-any.whl (3.3 MB)
  3.3/3.3 MB 12.3 MB/s eta 0:00:00
Installing collected packages: yt-dlp
Successfully installed yt-dlp-2025.5.22
YT-DLP no Google Colab - Download para MP3
-----
Cole a URL do vídeo que deseja baixar como MP3: 
```

Figura 9- Demonstração de funcionamento da Célula 6

6. DICAS ÚTEIS

Ao utilizar o Whisper no Google Colab, algumas práticas simples podem melhorar bastante sua experiência e evitar problemas comuns. Esta seção reúne orientações úteis para manter seu ambiente organizado, salvar corretamente seus arquivos e retomar o projeto quando necessário.

19

6.1. Salvando os arquivos gerados

Após a transcrição, os arquivos de saída (como .txt, .srt e .vtt) aparecem na aba lateral esquerda do Colab. Para baixá-los para o seu computador:

- Clique no ícone de **pasta** na lateral esquerda do Colab.
- Encontre o nome do arquivo que você deseja baixar.
- Clique com o **botão direito do mouse** sobre o nome do arquivo.
- Escolha a opção **Download**.

Faça isso antes de fechar o Colab, pois os arquivos são apagados ao encerrar a sessão.

6.2. Limpeza e organização da interface

Se você estiver reutilizando o notebook ou se deseja começar “do zero” visualmente, pode **limpar todas as saídas** e **recolher as células** para manter a interface limpa:

- No menu superior, clique em **Editar > Limpar todas as saídas**.
- Isso remove os resultados antigos e ajuda a evitar confusão.
- Você também pode clicar na **setinha lateral** ao lado dos títulos de cada seção para **recolher** as células e deixar a visualização mais enxuta.

6.3. Reutilizando o projeto

Para reutilizar o projeto em outro momento (mesmo dias depois), você pode abrir o notebook salvo no seu Google Drive:

- Acesse o Google Colab novamente.
- Clique em **Arquivo > Abrir > Google Drive**.
- Selecione o arquivo *Whisper_IdAD_5.0.ipynb*.
- Execute novamente as células na ordem correta.

O Colab salva automaticamente o notebook na pasta “**Colab Notebooks**” do seu Google Drive. Você pode renomear o notebook ou mover para outra pasta clicando com o botão direito sobre ele no Google Drive.

6.4. Recuperando o notebook original

Se você fizer alguma modificação acidental no código do notebook e ele parar de funcionar corretamente, não se preocupe. Você pode restaurar a versão original assim:

- Acesse o seu Google Drive e exclua a cópia modificada.
- Vá até o site do GP IdAD e baixe novamente o arquivo **.zip** atualizado.
- Extraia o arquivo e faça o upload do notebook original para o Colab, como descrito na seção 3 deste manual.

Isso garante que você sempre terá uma cópia limpa, funcionando corretamente, e evita a perda de tempo tentando corrigir erros inesperados.

7. ERROS COMUNS E SOLUÇÕES

21

Problema	Possível causa	Como resolver
Arquivo não encontrado (FileNotFoundException)	Você pode ter escrito o nome do arquivo errado ou esqueceu de enviá-lo para o Colab.	Confirme se o nome do arquivo está igual ao nome na lateral esquerda do Colab, incluindo '.mp3' ou '.mp4'. Ex: <code>entrevista.mp3</code>
O código parou ou travou	Às vezes a internet está lenta ou o Colab está sobrecarregado.	Espere um pouco. Se não funcionar, clique em 'Interromper execução' e depois em 'Executar novamente'.
Erro ao carregar o modelo (Model not found)	O nome do modelo pode estar digitado errado.	Use nomes exatos: "tiny", "base", "small", "medium" ou "large" (sempre entre aspas).
Transcrição vazia ou incompleta	O áudio pode estar muito baixo, com ruído ou em outro idioma.	Tente usar o modelo 'medium', escolha o idioma correto ou use a opção automática (<code>None</code>).
Colab se desconectou (Runtime disconnected)	O Colab ficou parado por muito tempo ou atingiu o tempo gratuito.	Clique em 'Conectar' no topo e depois execute as células de novo.
Texto sem acentos ou com erros	O modelo escolhido é muito simples.	Use um modelo melhor como o ' <code>medium</code> ' ou ' <code>large</code> ' para ter mais qualidade.
Erro com link de vídeo (yt-dlp)	O link do vídeo pode estar errado ou o vídeo é privado.	Verifique se o link do vídeo é válido e público. Tente abrir ele no navegador para testar.
Transcrição em idioma errado	Você escolheu o idioma errado.	Use a sigla certa (ex: ' <code>pt</code> ' para português, ' <code>en</code> ' para inglês) ou use <code>None</code> para detectar automático.
O código foi executado, mas não gerou nenhum arquivo	Pode ter dado erro silencioso ou o formato do arquivo está errado.	Tente usar arquivos <code>.mp3</code> , <code>.mp4</code> ou <code>.wav</code> . E veja se a célula deu alguma mensagem de erro embaixo.
O texto gerado está bagunçado	O áudio pode estar com muita fala sobreposta ou barulho.	Tente melhorar o áudio ou separar em partes menores. Verifique se o idioma está correto.

8. (ALGUMAS) CONSIDERAÇÕES FINAIS

22

Este guia foi desenvolvido com o objetivo de apoiar pesquisadores, estudantes e demais interessados no uso do **Whisper** de forma simples, gratuita e acessível. Por meio do Google Colab, é possível executar modelos avançados do Whisper, transformando arquivos de áudio e vídeo em texto com poucos cliques, sem a necessidade de instalações complexas ou equipamentos sofisticados.

Nosso objetivo é que este material sirva não apenas como um manual técnico, mas também como um instrumento de empoderamento digital, permitindo que qualquer pessoa, mesmo sem conhecimentos prévios em programação, consiga acessar e aplicar uma tecnologia de ponta em seus projetos acadêmicos, científicos, profissionais ou pessoais.

É fundamental compreender que, embora o Whisper apresente uma excelente capacidade de transcrição, ele não substitui o olhar crítico e atento do usuário. Questões como sotaques, ruídos de fundo, sobreposição de falas e qualidade do áudio podem impactar o resultado. Por isso, **a revisão humana continua sendo uma etapa indispensável** para garantir a precisão, a clareza e a fidedignidade dos materiais transcritos, especialmente em contextos de pesquisa, documentação ou divulgação científica.

Recomendamos que este guia seja explorado de forma dinâmica. Experimente diferentes modelos, teste arquivos variados, familiarize-se com as funções e adapte o uso da ferramenta às suas necessidades específicas. A tecnologia é uma aliada poderosa, e seu domínio abre portas para otimizar processos, agilizar análises e ampliar as possibilidades metodológicas em diversas áreas do conhecimento.

Esperamos que este material contribua para tornar seu trabalho mais ágil, produtivo e eficiente. Que ele facilite não apenas a transcrição, mas também incentive a exploração de novas tecnologias e ferramentas digitais no desenvolvimento de pesquisas e projetos.

9. NOTAS DO DESENVOLVEDOR

9.1. Nota sobre o desenvolvimento dos códigos

Os códigos que compõem este projeto foram desenvolvidos com o auxílio de ferramentas de inteligência artificial gerativa, em especial o modelo **DeepSeek**, que colaborou na redação, revisão e otimização dos scripts em Python utilizados no notebook desenvolvido. O uso de IA neste processo teve como finalidade ampliar a precisão do código, agilizar o desenvolvimento das funcionalidades e garantir maior clareza para o usuário final, sobretudo aqueles que não possuem familiaridade com programação. Ainda assim, todo o material foi cuidadosamente revisado, testado e adaptado para garantir sua aplicabilidade prática e sua adequação às finalidades deste guia.

9.2. Nota ao usuário: adaptação livre e criativa do código

O notebook disponibilizado neste projeto foi pensado para atender a uma ampla variedade de usos com foco na simplicidade e na acessibilidade. No entanto, reconhecemos que cada pesquisa tem suas particularidades — seja no tipo de material a ser transscrito, no idioma predominante, nos formatos de saída desejados ou nos fluxos de trabalho adotados.

Por isso, **incentivamos os usuários que tenham familiaridade com programação, ou que desejem experimentar, a modificar o código conforme suas necessidades específicas**. O código é aberto, comentado e organizado em blocos independentes, justamente para facilitar esse processo de adaptação. Seja ajustando parâmetros, automatizando etapas ou integrando com outras ferramentas, cada modificação pode tornar a aplicação ainda mais útil e alinhada ao seu contexto de pesquisa.

Explore, personalize e aproprie-se da ferramenta — ela é sua.

Dúvidas adicionais podem ser encaminhadas ao desenvolvedor responsável:

Thales Morbach Lange

Grupo de Pesquisa Ideologia e Análise de Discurso | wp.ufpel.edu.br/idad

Universidade Federal de Pelotas | portal.ufpel.edu.br

thales.lange@ufpel.edu.br | lattes.cnpq.br/3650755639394586

10. APÊNDICE: CHECKLIST PARA AUXÍLIO NO PASSO A PASSO DA EXECUÇÃO DO WHISPER

25

Etapa	Descrição / O que verificar	Feito ✓
PRÉ-REQUISITOS		
Conta Google ativa	Você precisa estar logado com sua conta do Google para acessar o Colab e o Drive.	
Conexão estável com a internet	Evita falhas no carregamento, instalação de pacotes e transcrição.	
CONFIGURAÇÃO INICIAL		
Notebook carregado no Colab	Entre em https://colab.research.google.com e clique em 'Upload' para enviar o arquivo .ipynb do projeto.	
EXECUÇÃO DO CÓDIGO		
Executar célula 1 (dependências)	Instala os pacotes necessários. Aguarde a mensagem de sucesso.	
Permitir acesso ao Google Drive	Quando aparecer mensagem de solicitação de acesso ao Google Drive, clique em 'Permitir'.	
Selecionar o modelo Whisper	Escolha um modelo (ex: 'medium') e execute a célula correspondente.	
Definir idioma do áudio	Configure com 'pt' para português ou None para detecção automática.	
Enviar e ajustar o nome do arquivo	Envie o arquivo no painel lateral ou indique o local do arquivo no Drive e indique o caminho correto na célula 4 ou 5, dependendo de qual for usar.	
Executar transcrição	Clique no ► da célula 4, ou a 5 se for processar vários arquivos. A transcrição iniciará.	
RESULTADOS E VERIFICAÇÃO		
Verificar mensagem de conclusão	Aguarde até aparecer 'Transcrição concluída com sucesso!'.	
Verificar arquivos gerados	Os arquivos devem aparecer na lateral. Caso tenha configurado para salvar no seu Drive, verifique no Google Drive se os arquivos foram salvos corretamente.	
Realizar download dos arquivos	Clique com botão direito > Download (ou, Baixar).	
Conferir a qualidade da transcrição	Leia atentamente o texto gerado e faça a revisão final.	