1stgeniusAI

GUIA GRÁTIS · OUVIR

GUIA GRATUITO · PASSO A PASSO

Faça a IA ouvir um vídeo, com o minuto de cada fala

Transforme o áudio de qualquer vídeo seu em texto com tempos, no seu computador, e peça ao Claude o gancho, a estrutura e as viradas.

PALAVRA-CHAVE

OUVIR · comente no Reel

ATUALIZADO

11/10/2026 · comandos testados por nós nesta data (macOS)

TRANSPARÊNCIA

Conteúdo produzido com apoio de IA. Avatar e voz dos vídeos são gerados por IA.

01 O que é e fonte oficial

O Claude lê texto e imagens, mas não assiste a vídeo: a página oficial de visão da Anthropic lista imagens (JPEG, PNG, GIF e WebP) e não lista vídeo. Se você cola o link de um vídeo e pede uma análise, o que vem de volta pode ser chute. O caminho que funciona é dar a ele o que ele lê: **a fala do vídeo em texto, com o minuto de cada frase**.

Para isso usamos três programas gratuitos que rodam no seu computador, sem enviar o áudio para ninguém:

- **whisper.cpp** (programa `whisper-cli`): transcreve a fala usando o modelo Whisper da OpenAI. Licença MIT.
- **ffmpeg**: extrai o áudio do vídeo. Licença GPL-3.0 ou posterior (como instalado pelo Homebrew).
- **yt-dlp** (opcional): baixa o áudio a partir de um link. Licença Unlicense (domínio público).

● FONTES OFICIAIS
github.com/ggml-org/whisper.cpp · github.com/yt-dlp/yt-dlp · platform.claude.com/docs/en/build-with-claude/vision.

CONFERIDO EM 11/10/2026

02 Pré-requisitos

- Mac com Homebrew (testado em Mac com chip Apple). Windows e Linux têm instruções no README do whisper.cpp; **não testamos** nesses sistemas.
- Espaço em disco: o modelo `small` pesa 466 MiB; o `large-v3-turbo`, 1,5 GiB (mais preciso).
- Um vídeo **seu** (ou de alguém que autorizou). Veja a seção 6.
- Acesso ao Claude (app, site ou Claude Code) para a etapa da análise.

03 Instalação passo a passo

1. Instale os programas (um comando só):

● TERMINAL

```
brew install whisper-cpp yt-dlp ffmpeg
```

2. Baixe um modelo para a pasta `~/.cache/whisper`. Para começar leve, o `small`:

TERMINAL

```
mkdir -p ~/.cache/whisper
curl -L -o ~/.cache/whisper/ggml-small.bin \
  https://huggingface.co/ggerganov/whisper.cpp/resolve/main/ggml-small.bin
```

Para mais precisão, troque por `ggml-large-v3-turbo.bin` (mesma pasta de download).

3. Salve o script abaixo como `video-para-texto.py`. Ele recebe um arquivo ou um link, extrai o áudio, transcreve e imprime uma linha por trecho com `[minuto:segundo]`:

SCRIPT · VIDEO-PARA-TEXTO.PY

```
#!/usr/bin/env python3
"""Vídeo (arquivo ou link) -> texto com o minuto de cada frase, tudo no seu
computador.
Uso: python3 video-para-texto.py <arquivo-ou-link> [modelo.bin]
Precisa de: ffmpeg, whisper-cli (whisper-cpp) e, para links, yt-dlp."""
import os, re, subprocess, sys, tempfile, time
alvo = sys.argv[1]
modelo = sys.argv[2] if len(sys.argv) > 2 else
os.path.expanduser("~/.cache/whisper/ggml-large-v3-turbo.bin")
tmp = tempfile.mkdtemp()
if alvo.startswith("http"):
    subprocess.run(["yt-dlp", "-q", "-x", "--audio-format", "wav", "-o", f"
{tmp}/baixado.%(ext)s", alvo], check=True)
    fonte = f"{tmp}/baixado.wav"
else:
    fonte = alvo
wav = f"{tmp}/audio16k.wav"
subprocess.run(["ffmpeg", "-y", "-loglevel", "error", "-i", fonte, "-vn", "-ac", "1",
"-ar", "16000", wav], check=True)
t0 = time.time()
r = subprocess.run(["whisper-cli", "-m", modelo, "-l", "pt", "-f", wav],
capture_output=True, text=True, check=True)
dur = time.time() - t0
for m in re.finditer(r"\[(\d+):(\d+):(\d+)\.\d+ --> [^\]]+\]\s+(.*)", r.stdout):
    h, mi, s, txt = m.groups()
    print(f"[{int(h)*60+int(mi):02d}:{int(s):02d}] {txt.strip()}")
print(f"# transcrito em {dur:.1f} s", file=sys.stderr)
```

04 Primeiros passos (menos de 10 minutos)

1. Com um arquivo de vídeo seu:

TERMINAL · ARQUIVO

```
python3 video-para-texto.py meu-video.mp4 ~/.cache/whisper/ggml-small.bin > texto.txt
```

2. Ou com o link de um vídeo público seu (precisa de internet):

TERMINAL · LINK

```
python3 video-para-texto.py https://www.instagram.com/reel/CODIGO/
~/.cache/whisper/ggml-small.bin > texto.txt
```

3. Abra `texto.txt` e confira. Este é o resultado real que obtivemos no nosso Reel 03 (20,8 s de vídeo):

RESULTADO REAL · REEL 03

```
[00:00] Seu anúncio pode estar gastando dinheiro à toa.
[00:02] Existe uma skill para o Cloud chamada Cloud Ads,
[00:05] feita para campanhas pagas no Google, Meta, YouTube, TikTok e outras
plataformas.
[00:10] Ela faz auditoria, monta estratégia e planeja orçamento e medição.
[00:14] E só mexe numa campanha com a sua aprovação explícita.
[00:17] Quer o passo a passo para testar? Comenta a guia aqui embaixo.
```

Medido em 11/10/2026 (Mac com chip Apple M3 Max): o vídeo de 20,8 s foi transcrito em 1,1 a 1,2 s com `large-v3-turbo` e em 0,9 s com `small`. Um áudio de 191 s (três Reels emendados três vezes) levou 5,9 s com `large-v3-turbo`. Não medimos vídeos de horas; o tempo cresce com a duração. A partir de um link, o teste com um Reel público nosso funcionou sem login; posts privados ou bloqueados podem falhar.

ATENÇÃO

Confira nomes próprios. Nos dois modelos o nome "Claude" saiu como "Cloud" na transcrição. Releia e corrija nomes de marcas, siglas e números antes de usar.

4. Cole o conteúdo de `texto.txt` no Claude com um dos prompts abaixo.

05 Prompts prontos para copiar

● PROMPT PRONTO · COPIE E COLE

Abaixo está a transcrição de um vídeo curto, com o minuto de cada trecho. Você não assistiu ao vídeo: trabalhe somente com este texto e diga quando algo não dá para saber só pela fala.

Me devolva:
1) o gancho (o que é dito nos primeiros 3 segundos) e por que ele prende ou não;
2) a estrutura em blocos, cada um com o minuto em que começa;
3) a virada principal (o minuto em que a ideia muda);
4) o pedido final (CTA) e o minuto;
5) 3 melhorias concretas para o meu próximo vídeo.

TRANSCRIÇÃO:
[cole aqui o texto]

● PROMPT PRONTO · COPIE E COLE

Esta é a transcrição de uma aula/podcast com o minuto de cada fala. Monte um índice: um título curto por assunto, o minuto em que o assunto começa e uma frase de resumo. Depois liste os 5 trechos mais fortes para virar vídeos curtos, com minuto inicial e final sugeridos. Se a transcrição parecer errada em algum ponto (nome próprio, número), marque com [conferir].

TRANSCRIÇÃO:
[cole aqui o texto]

● PROMPT PRONTO · COPIE E COLE

Aqui está a transcrição do MEU vídeo [tema]. Reescreva só os primeiros 2 segundos em 3 versões diferentes de gancho: uma que desmente uma crença, uma com número concreto e uma com pergunta. Mantenha o resto igual, não prometa dinheiro nem resultado garantido e use frases curtas, em português do Brasil.

TRANSCRIÇÃO:
[cole aqui o texto]

Exemplo real (resposta do Claude à transcrição do Reel 03, resumida)

RESUMO

Gancho [00:00]: pergunta-dor sobre gastar dinheiro à toa, sem apresentar o produto ainda.
Virada [00:02]: aparece a solução (uma skill para o Claude).
Prova [00:05 a 00:14]: lista de plataformas e de funções (auditoria, estratégia, orçamento, medição).
Limite de segurança [00:14]: só altera campanha com aprovação.
CTA [00:17]: pede o passo a passo por comentário.
Não dá para saber só pela fala: o que aparece na tela, o ritmo dos cortes e o visual.

06 Segurança, direitos e o que não faz

ATENÇÃO · SEGURANÇA

Use em vídeos seus, ou de quem autorizou, ou para estudo pessoal. Respeite os termos de cada plataforma e os direitos autorais; não republique o vídeo nem a transcrição de terceiros como se fossem seus.

- Funciona **só com o que é falado**. Imagem, música e texto na tela não entram na transcrição. Para o visual, use capturas de tela e envie como imagem ao Claude.
- Fala com ruído, sotaque forte ou vários falantes pode gerar erros. A transcrição não identifica quem está falando.
- O áudio é processado no seu computador; o que você cola no Claude vai para o serviço que você usa. Não cole conteúdo confidencial.
- O Claude pode errar ou inventar detalhes. Revise a análise antes de decidir qualquer coisa.
- Este guia não promete resultado de alcance ou de vendas.

07 Fontes

whisper.cpp (README e tabela de modelos): github.com/ggml-org/whisper.cpp · yt-dlp: github.com/yt-dlp/yt-dlp · Homebrew (licenças e versões: whisper.cpp 1.9.5, yt-dlp 2026.8.19, ffmpeg 9.0.2): formulae.brew.sh · Anthropic, documentação de visão (formatos de imagem aceitos): platform.claude.com/docs/en/build-with-claude/vision. Todas conferidas em 11/10/2026. Os tempos e a transcrição de exemplo foram medidos por nós nesta data. Guia educativo, sem garantia de resultado.

1stgeniusAI

PRÓXIMO PASSO · OUVIR

Da transcrição à sua próxima fala.

Escolha uma ideia do texto que você extraiu e escreva uma fala curta. O próximo vídeo mostra um caminho para trabalhar a voz.

Ver o próximo vídeo · VOZ

Quer o guia desse vídeo? Comente VOZ no Reel e confira a DM. O material é gratuito, sem cadastro de e-mail obrigatório.

Siga @1stgeniusai para acompanhar os próximos testes. Escolha apenas uma ideia para colocar em prática agora.

GUIAS GRATUITOS · SEM VENDA DE IMPLANTAÇÃO

O canal não oferece montagem de funil neste material. Pedido de guia não é autorização para campanhas ou newsletter. Nenhuma promessa de alcance, economia ou renda.

Fechamento revisado em 22/10/2026. As páginas técnicas anteriores mantêm suas datas e descrevem testes históricos; não foram revalidadas por esta atualização. Hoje o canal usa OpenReply, não ManyChat.

Conteúdo produzido com apoio de IA e revisado por pessoa.