



# WakkaQt

O MANUAL OFICIAL EXTRAOFICIAL

*"Porque Auto-Tune é caro e a acústica do seu banheiro só ajuda até certo ponto." Um guia de campo para ligar um microfone, se envergonhar com confiança e sair de lá com um vídeo suspeitosamente profissional.*

COBRE A VERSÃO 2.7.7

Sem assinaturas. Sem nuvem. Sem julgamento. (O monitor de afinação pode discordar.)

# Sumário

---

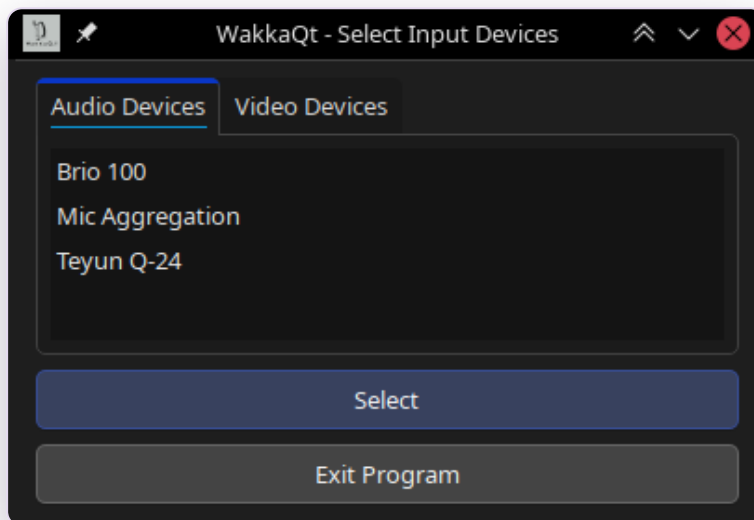
|    |                                                                |    |
|----|----------------------------------------------------------------|----|
| 01 | Escolhendo Suas Armas — a caixa de diálogo de dispositivos     | 4  |
| 02 | O Palco Principal — um tour pela janela que te julga           | 5  |
| 03 | Alimentando a Fera — carregando arquivos e saqueando o YouTube | 6  |
| 04 | Hora do Show — apertando o grande botão brilhante SING         | 8  |
| 05 | O Estúdio de Karaokê — afinação, sincronia e efeitos especiais | 9  |
| 06 | A Feiticeira da Base Instrumental — remoção de voz por IA      | 12 |
| 07 | Renderizando Sua Obra-Prima                                    | 13 |
| 08 | A Biblioteca de Sessões — sua vergonha, arquivada para sempre  | 14 |
| A  | Apêndice: A Lista Completa de Recursos                         | 15 |
| B  | Apêndice: Compilando no Linux (um recurso, não um defeito)     | 16 |

**O que é essa coisa?** WakkaQt é um estúdio de gravação e produção de karaokê gratuito e de código aberto, construído sobre o Qt6. Você carrega um vídeo de karaokê, canta em um microfone, e o software passa sua voz por um pipeline de processamento digital de sinais genuinamente sério — correção de afinação, redução de ruído, reverb, dinâmica — antes de renderizar um MP4 finalizado com sua webcam e um medidor de afinação que existe puramente para te manter honesto. Este manual percorre cada tela, na ordem em que você provavelmente vai realmente encontrá-las, com capturas de tela reais e uma quantidade nada recomendável de piadas.

# Escolhendo Suas Armas

*No qual você diz ao software o que ele tem permissão para ouvir.*

Toda sessão do WakkaQt começa com um interrogatório. Antes de você ter permissão para chegar perto da janela principal, o aplicativo quer saber exatamente qual microfone e (opcionalmente) qual webcam ele está autorizado a usar. Isso é bom: significa que o WakkaQt nunca está adivinhando, e nunca está gravando silenciosamente do fone USB errado enquanto você serenata um dispositivo que nem está conectado.



*A caixa de diálogo Select Input Devices. Duas abas — Audio Devices e Video Devices — listam tudo que o seu sistema operacional atualmente admite possuir.*

- 1 Clique na aba **Audio Devices** e escolha seu microfone na lista. O WakkaQt enumera todo dispositivo de entrada que o seu SO reporta, então se você tem um microfone condensador USB, uma configuração de múltiplos microfones agregados, ou só a batatinha embutida do laptop, vai aparecer aqui.
- 2 Mude para a aba **Video Devices** se quiser imagens de webcam na renderização final. Esse passo é totalmente opcional — pule e o WakkaQt roda feliz apenas com áudio, sem precisar de câmera, sem drama.
- 3 Pressione **Select** para confirmar e abrir a janela principal. Ou pressione **Exit Program** se você mudou de ideia sobre essa história toda de "cantar". Sem julgamento. Por enquanto.

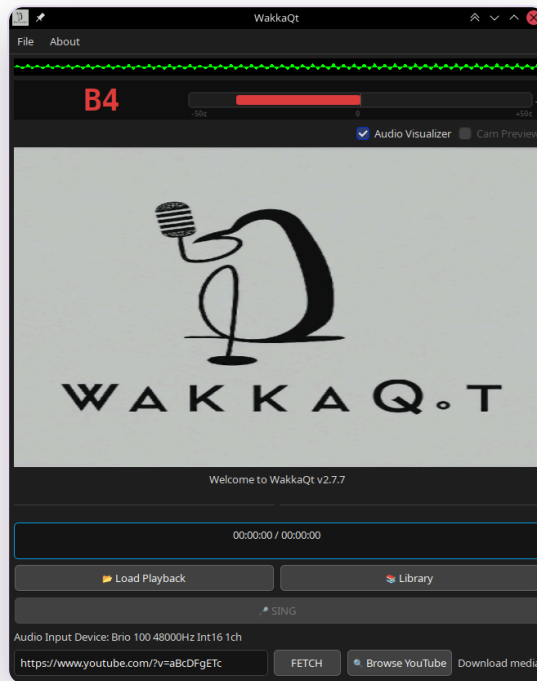
## DICA DE MESTRE

Não conectou uma webcam? Sem pânico, e não desconecte nada às pressas — o WakkaQt recorre a um pipeline limpo, só de áudio, do início ao fim: gravação, pré-visualização e renderização pulam a câmera com elegância. Ninguém precisa te ver assim.

# O Palco Principal

*Um tour guiado pela janela que você vai encarar pelas próximas horas.*

Esta é a base de operações. Toda gravação, toda renderização e todo arrependimento começam e terminam aqui. Vamos ver o que realmente está na tela, de cima a baixo.



*A janela principal em repouso. Calma. Paciente. Esperando.*

- **Monitor de afinação (faixa superior):** sempre visível, sempre ouvindo, sempre pronto para exibir uma letra enorme tipo **B4** ao lado de uma barra de cents que fica verde, amarela ou vermelha dependendo do quão perto você está da nota que deveria estar cantando. Fica ligado mesmo quando você não está gravando — feedback grátis e não solicitado.
- **Alternância Audio Visualizer:** silencia ou reativa as duas barras de forma de onda ao lado da pré-visualização da webcam — elas mostram o áudio da própria reprodução, não o seu microfone.
- **Área de vídeo:** mostra o vídeo de karaokê carregado, ou o logo do WakkaQt se nada foi carregado ainda, esperando pacientemente que você alimente com alguma coisa.
- **Barra de transporte / status:** tempo decorrido, título da música atual quando carregada, e os controles de reprodução assim que há mídia pronta.
- Botões **Load Playback** e **Library**: abrem um arquivo do disco, ou reabrem uma sessão anterior do seu acervo pessoal cada vez maior de tentativas de karaokê.
- Botão **SING**: grande, impossível de não ver, levemente intimidador. Coberto com detalhes no Capítulo 4.
- Linha **Audio Input Device**: um lembrete permanente de qual microfone você escolheu lá no Capítulo 1, incluindo a taxa de amostragem e profundidade de bits ao vivo, para quem curte esse tipo de detalhe.
- **Campo de URL + FETCH / Browse YouTube / Download media:** o caminho mais rápido entre "eu quero cantar essa música" e "eu já estou cantando essa música". Mais no Capítulo 3.

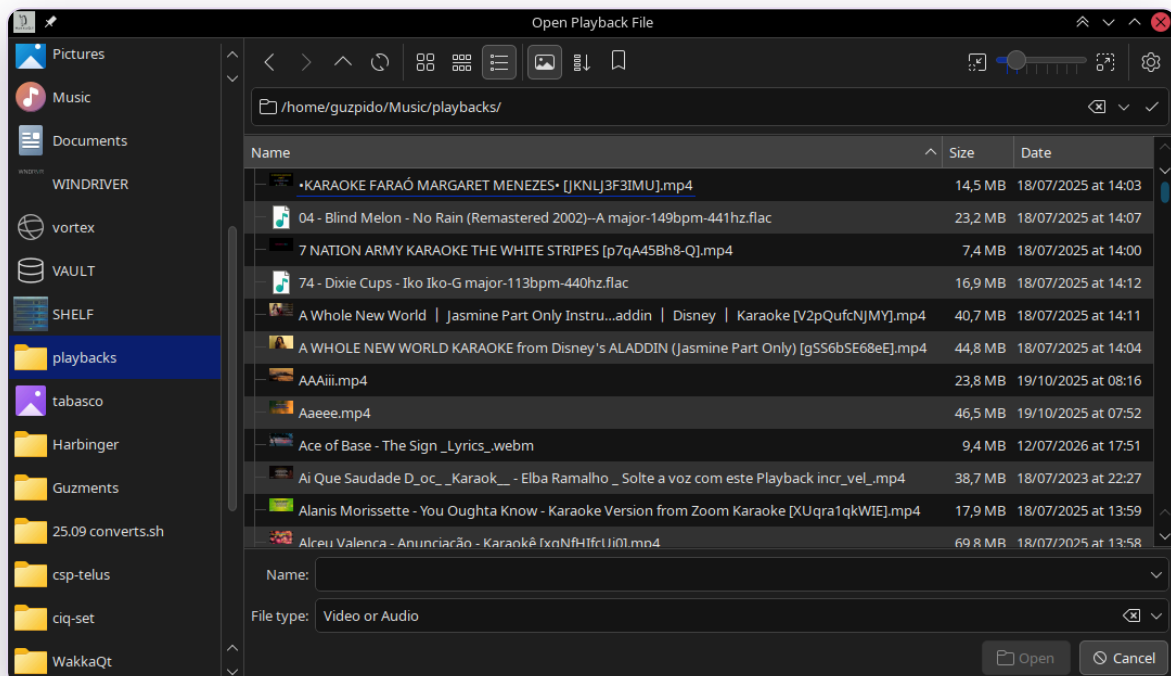
# Alimentando a Fera

*Colocando uma música na tela, por qualquer meio necessário.*

O WakkaQt não se importa de onde vem a sua faixa de karaokê, contanto que seja um formato que o Qt6 Multimedia entenda: MP4, MKV, WebM, AVI, MOV, MP3, WAV, FLAC ou OPUS. Existem duas formas respeitáveis de carregar uma, e uma terceira deliciosamente preguiçosa.

## Opção A — Abrir um arquivo que você já tem

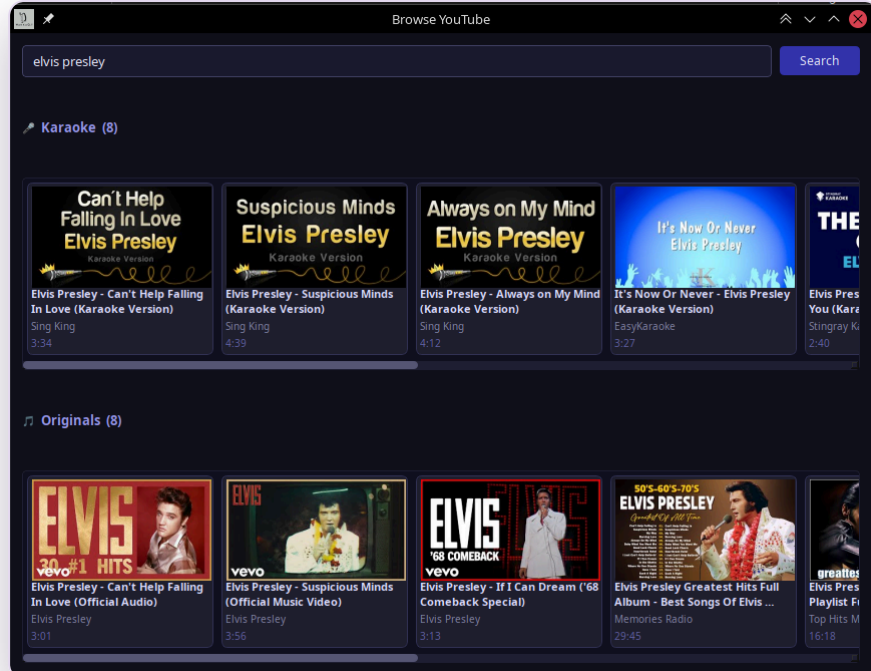
Clique em **Load Playback** e escolha um arquivo da sua coleção existente. Este é o seletor de arquivos padrão, sem surpresas, que o seu sistema operacional já oferece — nada a explicar aqui, exceto que, a julgar pela captura de tela abaixo, algumas pessoas já têm *muchos* arquivos de karaokê.



*A caixa de diálogo Open Playback File. Sim, essa é uma pasta cheia de centenas de megabytes dos maiores sucessos de outras pessoas, e não, não vamos falar sobre isso.*

## Opção B — Navegar pelo YouTube de dentro do aplicativo

Clique em **Browse YouTube** para abrir uma caixa de diálogo de busca de verdade, com duas linhas roláveis de resultados — versões dedicadas de **Karaokê** em cima, **Originais** embaixo, para o caso de você preferir aprender a música antes de se humilhar com a base instrumental. Clique num card para pré-visualizar, depois em "Download this video" para mandar direto para o fluxo de download (movido a `yt-dlp` por baixo dos panos).



Buscar por "elvis presley" traz oito versões de karaokê e oito originais. Escolha ambiciosa para uma terça-feira.

### OPÇÃO C — O JEITO PREGUIÇOSO

Já tem um link do YouTube na área de transferência? Cole direto no campo de URL da janela principal e aperte **FETCH**. O mesmo pipeline de download via `yt-dlp`, sem precisar navegar nada.

Assim que algo é carregado, a janela principal ganha vida: o vídeo toca, uma forma de onda rola na tela, e a barra de transporte completa (retroceder, play/pause, parar, avançar) aparece acima dos botões.



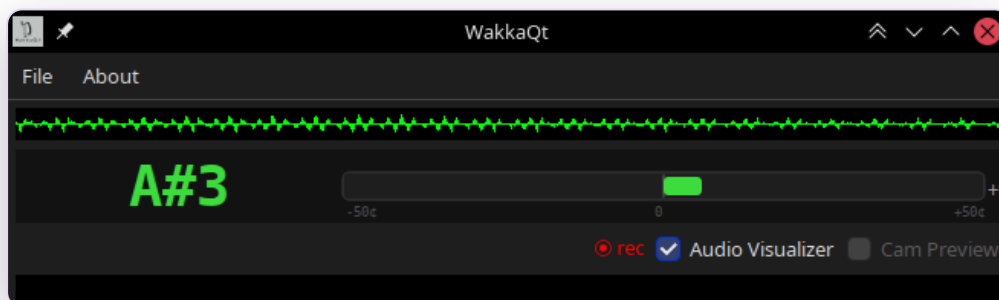
Uma faixa carregada e tocando — capa do álbum, título, e uma forma de onda ao vivo, com Load Playback, Load Again, Library, SING e Backing Track todos à disposição.

# Hora do Show

*O botão para o qual você vem se preparando. Aperte-o. Nós desafiaremos você.*

Com uma música carregada e um dispositivo selecionado lá no Capítulo 1, o botão  **SING** ganha vida. Aperte-o, e é exatamente isto que acontece:

- 1 A reprodução começa do início (ou de onde você deixou pausado por último).
- 2 Seu microfone começa a gravar, ao vivo, sem compressão, na sua taxa de amostragem nativa — sem downsampling forçado, então se o seu microfone faz 48 kHz, o WakkaQt processa em 48 kHz do começo ao fim.
- 3 Se você ativou uma webcam lá no Capítulo 1, ela grava junto com o áudio, sincronizada ao mesmo relógio.
- 4 A faixa do monitor de afinação — aquela que ficou silenciosamente te observando esse tempo todo — muda para o modo de gravação: um indicador vermelho de **rec** aparece, e o nome da nota alterna entre verde, amarelo e vermelho em tempo real enquanto você persegue (ou foge d) a afinação correta.
- 5 A gravação para automaticamente quando a música termina, ou você pode parar manualmente a qualquer momento.



*No meio da gravação: o indicador rec está aceso, a nota marca A#3, e a barra está confortavelmente no verde. É assim que fica quando está indo bem.*

## UM PÉ NO CHÃO

Nada está sendo auto-tunado ao vivo. O monitor de afinação é puramente informativo durante a gravação — um treinador parado atrás de você com uma prancheta. A correção de verdade acontece depois, no Estúdio de Karaokê, onde você tem controle total sobre quanto (ou quão pouco) aplicar dela.

Assim que a música termina e a gravação para, o WakkaQt abre automaticamente o **Estúdio de Karaokê** — a caixa de diálogo de pré-visualização e aprimoramento — com sua voz recém-gravada já carregada e esperando.

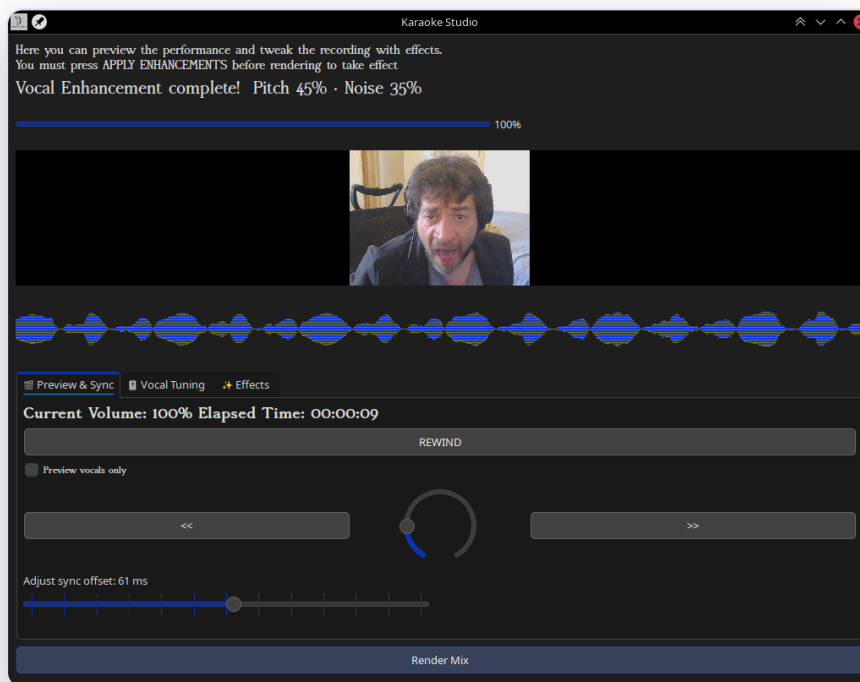
# O Estúdio de Karaokê

*Onde esforço humano bruto silenciosamente vira algo que você assumiria nas redes sociais.*

O Estúdio de Karaokê (tecnicamente a caixa de diálogo Preview, mas "Karaoke Studio" é como ela orgulhosamente se autodenomina na barra de título) é onde a sua gravação vocal é processada, pré-visualizada, ajustada, e pré-visualizada de novo — quantas vezes sua paciência permitir — antes de você se comprometer com uma renderização final. Está organizado em três abas, e nenhuma delas é opcional de entender.

## Aba 1 — Preview & Sync

A aba padrão. Uma pré-visualização ao vivo da webcam (se você gravou uma) fica acima de um visualizador de áudio rolante, mantido perfeitamente sincronizado com a reprodução — retroceda, avance ou ajuste o deslocamento, e o vídeo ressincroniza imediatamente. Abaixo disso: controles de transporte, uma caixa de seleção "Preview vocals only" para isolar sua voz da base instrumental, botões de busca, e — crucialmente — o controle deslizante de **sync offset**, para os inevitáveis poucos milissegundos de deriva entre sua voz e o vídeo.



*Preview & Sync — aprimoramento já aplicado (Pitch 45% · Noise 35%), deslocamento de sincronia ajustado em 61 ms.*

## Aba 2 — Vocal Tuning

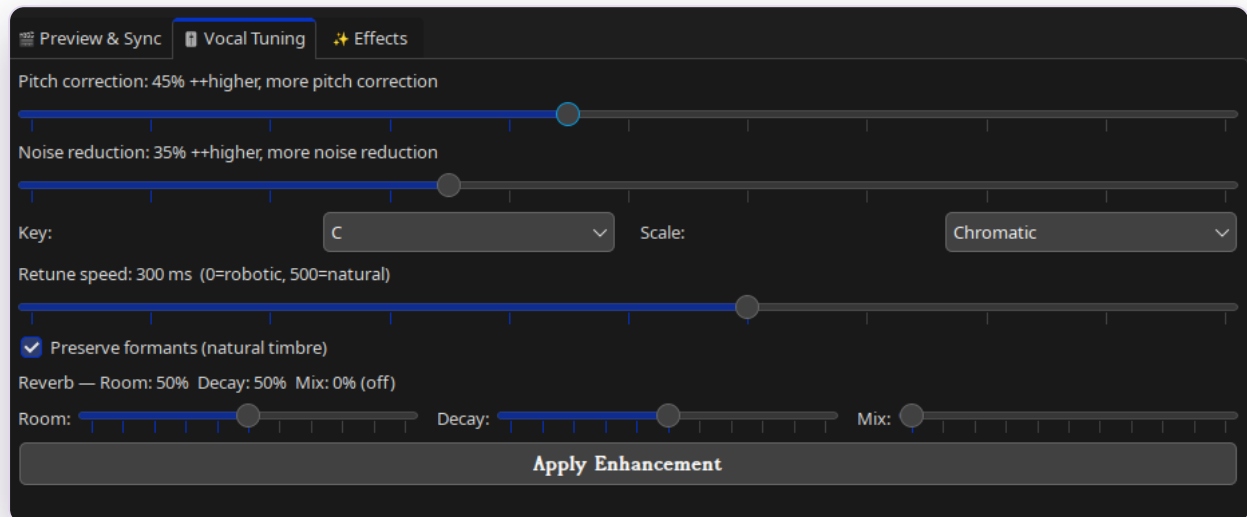
Este é o pipeline de processamento de sinais de verdade, exposto como controles deslizantes em vez de jargão intimidador de engenheiro de áudio:

- **Pitch correction (correção de afinação)** — mudança de tom via phase-vocoder. 0% deixa sua voz completamente em paz; valores mais altos encaixam com mais força na nota-alvo.
- **Noise reduction (redução de ruído)** — um gate de subtração espectral com estimativa adaptativa do piso de ruído, para expulsar o zumbido do ventilador, o barulho do teclado e o soprador de folhas do vizinho.
- **Key & Scale (tom e escala)** — escolha um tom (qualquer um dos 12) e uma escala (Maior, Menor, Pentatônica, Blues, Dórica, Mixolídia, Lídia, Frígia, Lócria, Menor Harmônica, Menor Melódica, Tons Inteiros, Diminuta, ou simplesmente Cromática se preferir não se comprometer) para que a correção de afinação só encaixe em notas que realmente pertencem à música.



- **Retune speed (velocidade de reafinação)** — o ms para o inconfundível efeito robótico total estilo T-Pain; até 500 ms para um deslize lento e natural que mal admite estar corrigindo alguma coisa.
- **Preserve formants (preservar formantes)** — uma etapa de ressíntese de envelope baseada em LPC que mantém sua voz soando como a de um ser humano mesmo depois de uma mudança de tom agressiva, em vez de um esquilo com opiniões.
- **Reverb** — reverb estilo Schroeder ao estilo Freeverb, com controles independentes de Room (ambiente), Decay (decaimento) e Mix, para transformar seu quarto numa catedral, ou pelo menos num quarto um pouco maior.

Ajuste o que quiser, pressione **Apply Enhancement**, e a pré-visualização atualiza para você ouvir exatamente o que fez consigo mesmo.



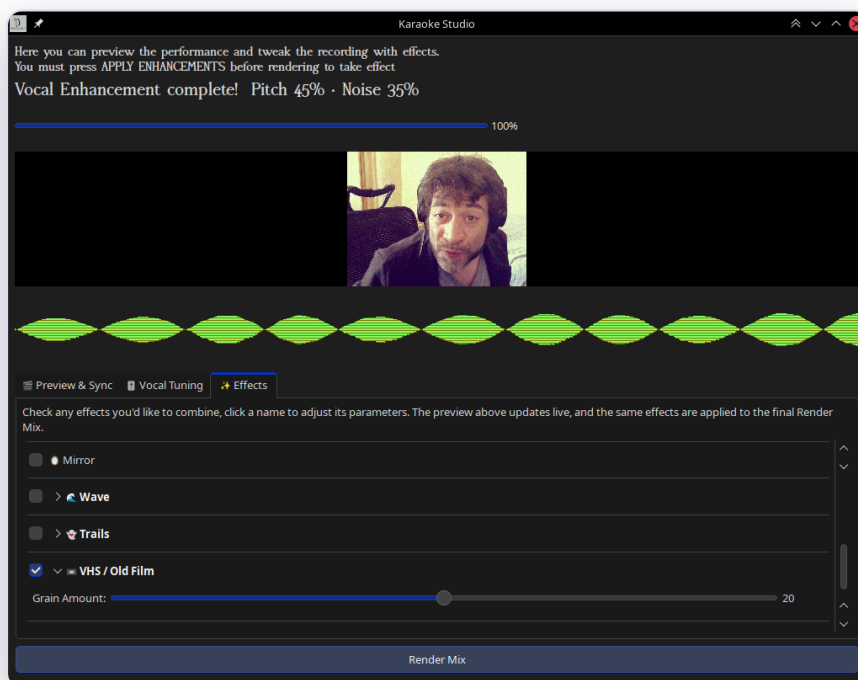
*Vocal Tuning: correção de afinação em 45%, redução de ruído em 35%, tom de C, escala Cromática, velocidade de reafinação 300 ms, formantes preservados, reverb parado em 0% de mix (ou seja, desligado, por enquanto).*

# Efeitos Especiais

*Porque uma afinação tecnicamente correta não é nem de longe espetáculo suficiente.*

## Aba 3 — Effects

Doze efeitos de vídeo combináveis, cada um representado por um ou dois controles deslizantes amigáveis no lugar de uma cadeia bruta de `libavfilter` que você nunca vai precisar digitar na mão:  Vertigo,  Technicolor,  Sépia,  Alto Contraste,  Preto e Branco,  Vinheta,  Negativo,  Espelho,  Onda,  Rastros,  VHS/Filme Antigo, e  Detecção de Bordas. Marque quantos quiser — eles se encadeiam na ordem escolhida — e a pré-visualização ao vivo acima atualiza instantaneamente, usando exatamente a mesma cadeia de filtros que vai para a renderização final. O que você vê é realmente o que você leva.



*A aba Effects, com VHS / Old Film marcado e seu controle deslizante Grain Amount no 20 — porque se você vai reviver a sua juventude, também pode parecer uma fita de camcorder dela.*

### CONTEÚDO BÔNUS

 Vertigo roda sobre um plugin `frei0r` em vez de um filtro comum. Se esse plugin não estiver instalado no seu sistema, o WakkaQt simplesmente esconde a opção da lista em vez de te jogar um erro na cara — sem crash, sem drama, só uma forma a menos de fazer sua webcam girar ameaçadoramente.

Assim que tudo estiver soando e parecendo certo, pressione **Render Mix** na parte de baixo da caixa de diálogo para seguir para o Capítulo 7. Ou continue ajustando. O Estúdio vai esperar. Ele não tem outro lugar para estar.

# A Feitiçaria da Base Instrumental

*Transforme qualquer música numa faixa de karaokê, usando uma pequena IA que é toda sua.*

Carregou uma música normal — com voz e tudo, sem etiqueta de "instrumental" ou "karaokê" à vista — e quer cantar em cima dela mesmo assim? Clique em  **Backing Track**, visível logo abaixo do botão SING sempre que um arquivo está carregado.

- 1 No primeiro uso, o WakkaQt baixa o modelo ONNX de separação vocal **UVR-MDX-NET-Inst\_HQ\_3** (cerca de 80 MB), verifica contra um checksum fixo, e guarda em cache em `~/WakkaQt/models/`. Isso acontece exatamente uma vez, para sempre.
- 2 O modelo separa a voz do instrumental usando deep learning MDX-Net, rodando inteiramente na sua própria máquina através de um pipeline completo de STFT/iSTFT. Sem necessidade de conexão à internet depois desse primeiro download, sem serviço na nuvem envolvido, nada sai do seu computador.
- 3 Se a origem era um arquivo de vídeo, o instrumental separado é remixado direto de volta no vídeo original por padrão, para você manter as imagens. Prefere só áudio? Escolha WAV ou MP3 na caixa de diálogo de salvar.
- 4 Mudou de ideia no meio do caminho? O botão **Abort** cancela a separação de forma limpa, em pleno voo, sem deixar saída corrompida para trás.

## POR QUE ISSO IMPORTA

Essa é toda a filosofia "sem assinatura, sem nuvem" resumida em um único recurso. Muitas ferramentas cobram mensalmente, ou enviam seu áudio para um servidor em algum lugar, para fazer exatamente isso. O WakkaQt só baixa um modelo pequeno uma vez e faz a conta localmente, na sua CPU, de graça, para sempre.

# Renderizando Sua Obra-Prima

*O momento em que sua performance com acústica de banheiro vira um arquivo permanente e compartilhável.*

Pressione **Render Mix** no Estúdio de Karaokê, e o WakkaQt monta tudo em um MP4 finalizado de 1920×1080:

- Só a imagem da sua webcam aparece no vídeo final — o vídeo de karaokê nunca é incluído, só seu áudio.
- Sua voz é mixada com todo aprimoramento do Capítulo 5 totalmente aplicado — correção de afinação, redução de ruído, preservação de formantes, reverb, mais uma passada final de polimento com compressão, limitação e excitador harmônico para um master alto e limpo.
- Uma **faixa indicadora de afinação** é gravada diretamente no quadro da webcam: o nome da nota, uma barra de desvio em cents, e um valor `+/-Nc`, colorido em verde (afinado), amarelo (desviando), ou vermelho (passando por um momento) — um registro permanente e revisável de exatamente como sua afinação se comportou.
- A renderização roda nativamente em processo via as bibliotecas do FFmpeg com uma barra de progresso em tempo real, e tenta automaticamente codificação H.264 acelerada por hardware (NVENC, depois VAAPI, depois V4L2M2M) antes de recorrer à codificação por software — o que for mais rápido na sua máquina.
- Mudou de ideia no meio da renderização? **Abort Render** a interrompe de forma limpa, sem deixar um arquivo quebrado para trás.

## SEM BIBLIOTECAS DE DESENVOLVIMENTO DO FFMPEG NA HORA DE COMPILAR?

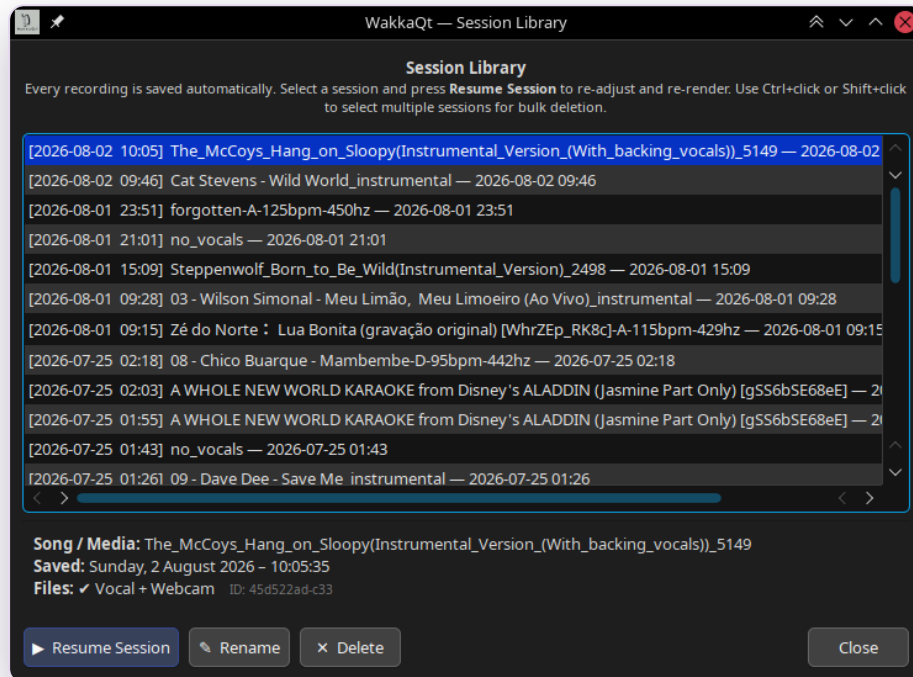
Sem problema. O WakkaQt recorre graciosamente a chamar um processo `ffmpeg` do sistema em vez disso. Um pouco menos elegante por baixo dos panos, resultado idêntico na tela.

Quando terminar, você tem um vídeo finalizado e compartilhável — e, quer você tenha pedido ou não, um registro permanente de cada nota que você tentou passar despercebida pelo monitor de afinação.

# A Biblioteca de Sessões

*Toda gravação, arquivada automaticamente, goste você ou não.*

Toda vez que você grava, o WakkaQt silenciosamente salva a sessão inteira em `~/WakkaQt/library/` — uma pasta única, todos os arquivos de origem, e um manifesto JSON descrevendo o que há dentro. Clique em **Library** na janela principal para ver o histórico completo e com data e hora de toda a sua carreira no karaokê.



A Session Library. Um registro contínuo, com data e hora, de toda música que você já cantou neste programa, remontando semanas. Não existe um botão "apagar histórico e fingir que isso nunca aconteceu", mas existe um botão Delete perfeitamente funcional para sessões individuais.

- **Resume Session** — recarrega qualquer gravação passada direto de volta no Estúdio de Karaokê, com todos os arquivos originais de voz e webcam, pronta para afinar e re-renderizar com configurações de aprimoramento completamente diferentes. A voz processada intermediária nunca é armazenada — ela é regenerada do zero a partir da sua gravação bruta toda vez, então a pré-visualização sempre reflete seus controles atuais, não os da semana passada.
- **Rename** — dê à sessão um nome que você realmente vai reconhecer depois, em vez do nome de arquivo original que calhou de ser.
- **Delete** — remove uma sessão para sempre. Segure **Ctrl** ou **Shift** enquanto clica para selecionar e apagar várias de uma vez, para aquela faxina geral ocasional.

## SALVO DE FORMA TRANSACIONAL

As sessões são escritas primeiro numa pasta temporária `.partial` e só renomeadas para o lugar definitivo depois que tudo foi salvo com sucesso. Uma queda do sistema ou uma tampa fechada no meio do salvamento não consegue deixar uma entrada quebrada e pela metade bagunçando sua biblioteca.

# A Lista Completa de Recursos

Para quem lê por cima, e para quem só quer as provas.

| Recurso                                                             | Status               |
|---------------------------------------------------------------------|----------------------|
| Reprodução de MP4 / MKV / WebM / MP3 / WAV / FLAC                   | ✓ Sim                |
| Gravação por microfone (dispositivo selecionável)                   | ✓ Sim                |
| Gravação por webcam                                                 | ✓ Sim                |
| Gravação somente de áudio (sem webcam obrigatória)                  | ✓ Sim                |
| Monitor de afinação em tempo real (algoritmo YIN, sempre ativo)     | ✓ Sim                |
| Visualizador de forma de onda em tempo real                         | ✓ Sim                |
| Correção de afinação (phase vocoder)                                | ✓ Sim                |
| Encaixe de afinação sensível a escala/tom                           | ✓ Sim                |
| Preservação de formantes (LPC)                                      | ✓ Sim                |
| Redução de ruído (subtração espectral)                              | ✓ Sim                |
| Reverb (Freeverb / Schroeder)                                       | ✓ Sim                |
| Compressor + limitador + excitador harmônico                        | ✓ Sim                |
| Caixa de pré-visualização com ajuste ao vivo                        | ✓ Sim                |
| Renderização nativa via FFmpeg (em processo)                        | ✓ Sim                |
| Sobreposição de afinação gravada no vídeo renderizado               | ✓ Sim                |
| Pré-visualização de vídeo da webcam ao vivo (sincronizada)          | ✓ Sim                |
| 12 efeitos de vídeo combináveis                                     | ✓ Sim                |
| Biblioteca de sessões (salvar / renomear / excluir / re-renderizar) | ✓ Sim                |
| Download do YouTube (via yt-dlp)                                    | ✓ Sim                |
| Navegador de karaokê do YouTube (busca + pré-visualização)          | ✓ Sim                |
| Separação vocal por IA → base instrumental (modelo ONNX local)      | ✓ Sim                |
| Base instrumental: saída que preserva o vídeo                       | ✓ Sim                |
| H.264 acelerado por hardware (VAAPI / NVENC)                        | ✓ Sim                |
| Abortar renderização / abortar separação                            | ✓ Sim                |
| Multiplataforma (Linux / Windows)                                   | ✓ Sim                |
| Assinatura obrigatória                                              | ✗ Não                |
| Envia dados para um servidor                                        | ✗ Não                |
| Julgamento sobre o seu canto                                        | principalmente ✗ Não |

# Compilando no Linux

*"Um recurso, não um defeito" — a posição oficial do projeto.*

Usuários de Windows ganham um binário pronto. Usuários de Linux ganham a construção de caráter. Aqui vai a versão resumida — veja o README do projeto para o passo a passo completo, distro por distro.

## Pré-requisitos no Debian / Ubuntu

```
sudo apt install build-essential cmake ninja-build qt6-base-dev qt6-multimedia-dev libfftw3-dev  
libavformat-dev libavcodec-dev libavfilter-dev libavutil-dev libswresample-dev libswscale-dev  
libglib2.0-dev pkg-config
```

## Configurar e compilar

```
cmake -B build -DCMAKE_BUILD_TYPE=Release  
cmake --build build --parallel  
./build/WakkaQt
```

### FORTEMENTE RECOMENDADO

Instale o `libavcodec-extra` (Debian/Ubuntu) para que as gravações realmente usem H264 em vez de recorrer ao MotionJPEG. E mantenha o `ffmpeg` e o `yt-dlp` no seu `$PATH` em tempo de execução — ambos são usados mesmo na compilação com FFmpeg nativo, como respaldo e para downloads, respectivamente.

### OPCIONAL: O RECURSO DE BASE INSTRUMENTAL POR IA

Instale o ONNX Runtime ( `libonnxruntime-dev` no Debian/Ubuntu) antes de configurar. É totalmente opcional — sem ele, o WakkaQt compila e roda normalmente, e o botão Backing Track simplesmente não aparece.

Instruções completas, nomes de pacotes para Fedora/RHEL, configuração do ONNX Runtime no Windows, e os passos de instalação do plugin freir para o efeito Vertigo estão todos no `readme.md` do projeto.

# Você Chegou ao Fim

*O que, estatisticamente, te coloca à frente da maioria das pessoas que abrem um manual de software. Vá carregar uma música, conecte um microfone, e aperte o grande botão brilhante. O monitor de afinação está esperando, e ele não julga.*

*(Ele julga um pouquinho.)*

A+

CERTIFICADO DE CONCLUSÃO DO MANUAL