VOICE2TEXT AI

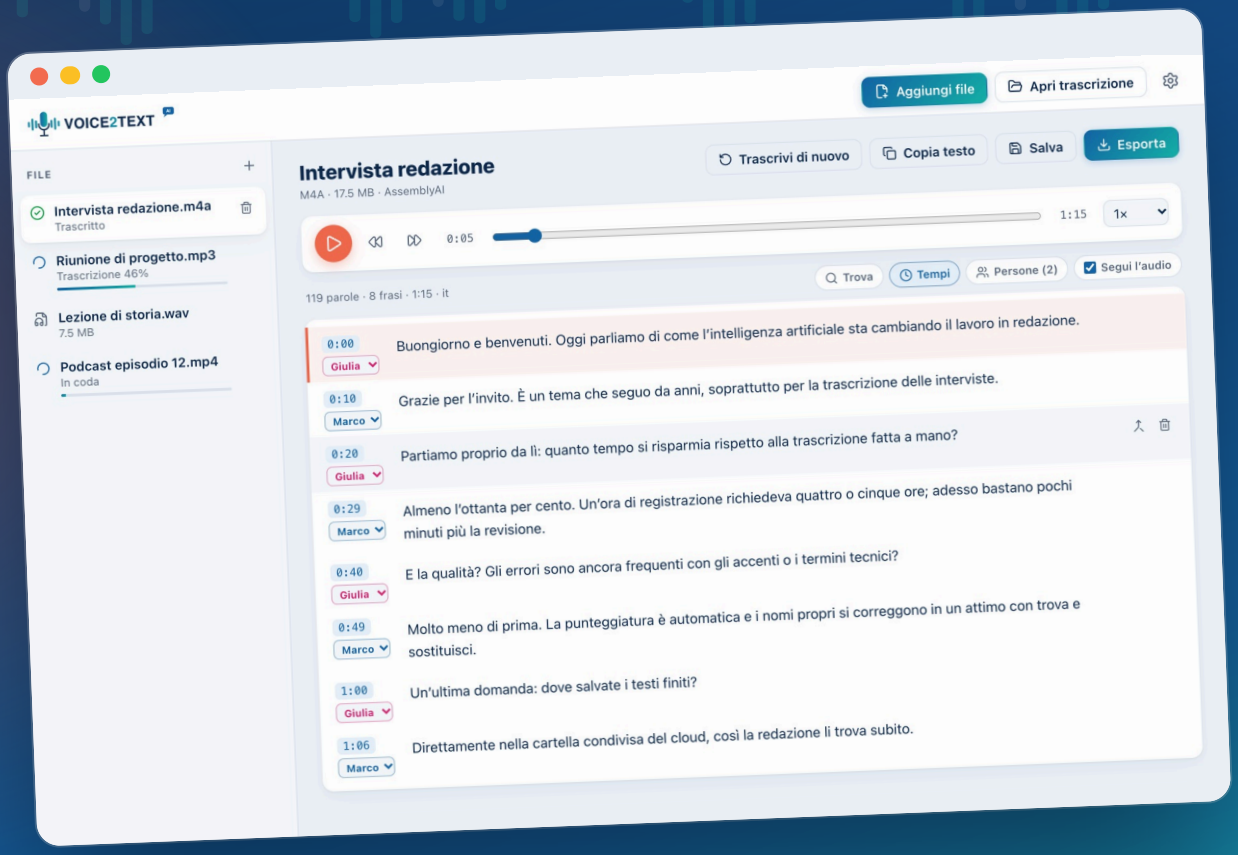
MANUALE D'USO

Voice2Text

Dalla voce al testo, in ogni lingua.

Trascrivi file audio e video con l'intelligenza artificiale, correggi il testo e salvalo dove vuoi.

Whisper offline · 12 formati · Cloud · 9 lingue



Versione 1.1 · macOS · Windows · Linux · Italiano

Indice

01	Benvenuto in Voice2Text	3
02	Installazione	4
03	La finestra principale	5
04	Scegliere motore e lingua	6
05	Servizi online: il tuo audio esce dal computer	8
06	Durante la trascrizione	10
07	Correggere la trascrizione	11
08	Salvare ed esportare	13
09	Salvare nel cloud	15
10	Impostazioni e lingue	17
11	Scorciatoie da tastiera	19
12	Domande frequenti	20

01 Benvenuto in Voice2Text

> Voice2Text trasforma registrazioni audio e video in testo scritto. Interviste, riunioni, lezioni, podcast, messaggi vocali: trascini il file nella finestra, scegli il motore di intelligenza artificiale e in pochi minuti hai il testo, pronto da correggere e da salvare.

Cosa puoi fare

- **Qualsiasi formato.** MP3, M4A, WAV, FLAC, OGG, AAC, AIFF, video MP4, MOV, MKV, WebM e molti altri; con ffmpeg installato praticamente tutti.
- **Anche senza Internet.** Whisper funziona direttamente sul computer: l’audio non lascia mai la tua scrivania.
- **Chi parla.** Con AssemblyAI e Deepgram le voci vengono separate e puoi dare un nome a ogni persona.
- **Correzione comoda.** Il testo è diviso in frasi collegate all’audio: un clic e riascolti il punto esatto.
- **Ogni formato in uscita.** Word, PDF, ODT, RTF, TXT, Markdown, HTML, sottotitoli SRT e VTT, CSV, JSON e il progetto riapribile.
- **Direttamente nel cloud.** Google Drive, OneDrive, Dropbox, Box, Proton Drive, iCloud e altri, tramite le loro cartelle sul computer.
- **Nove lingue.** L’interfaccia parla italiano, inglese, francese, spagnolo, tedesco, cinese, giapponese, russo e arabo; le traduzioni sono file XML che puoi correggere.

Requisiti

macOS 11 o successivo (Mac Apple o Intel), Windows 10/11 a 64 bit, oppure Linux a 64 bit (Ubuntu, Debian, Fedora e derivate). Per Whisper offline conviene avere almeno 8 GB di memoria; i servizi online richiedono solo una connessione.

02 Installazione

> Scarica il file per il tuo sistema e installalo come qualsiasi altro programma.

macOS

Apri `Voice2Text_1.1.1_universal.dmg` e trascina l'icona di Voice2Text nella cartella **Applicazioni**. Al primo avvio macOS potrebbe dire che lo sviluppatore non è verificato: apri **Impostazioni di Sistema** → **Privacy e sicurezza** e scegli **Apri comunque**, oppure fai clic destro sull'app e scegli **Apri**. Serve una volta sola.

Windows

Avvia `Voice2Text_1.1.1_x64-setup.exe`. Se compare la finestra blu di Windows SmartScreen, fai clic su **Ulteriori informazioni** e poi su **Esegui comunque**. L'installazione non richiede i permessi di amministratore.

Linux

Su Ubuntu e Debian installa il pacchetto `.deb` (doppio clic oppure `sudo apt install ./Voice2Text_1.1.1_amd64.deb`); su Fedora il pacchetto `.rpm`. In alternativa rendi eseguibile l'`.AppImage` (`chmod +x`) e avviala.

Facoltativo: ffmpeg. Voice2Text legge da solo i formati più diffusi. Per quelli rari (OPUS, WMA, AMR, AVI…) installa ffmpeg: `brew install ffmpeg` su Mac, `winget install ffmpeg` su Windows, `sudo apt install ffmpeg` su Linux. Viene trovato automaticamente.

03 La finestra principale

> Tutto il lavoro si svolge in un'unica finestra: a sinistra l'elenco dei file, al centro il file selezionato.

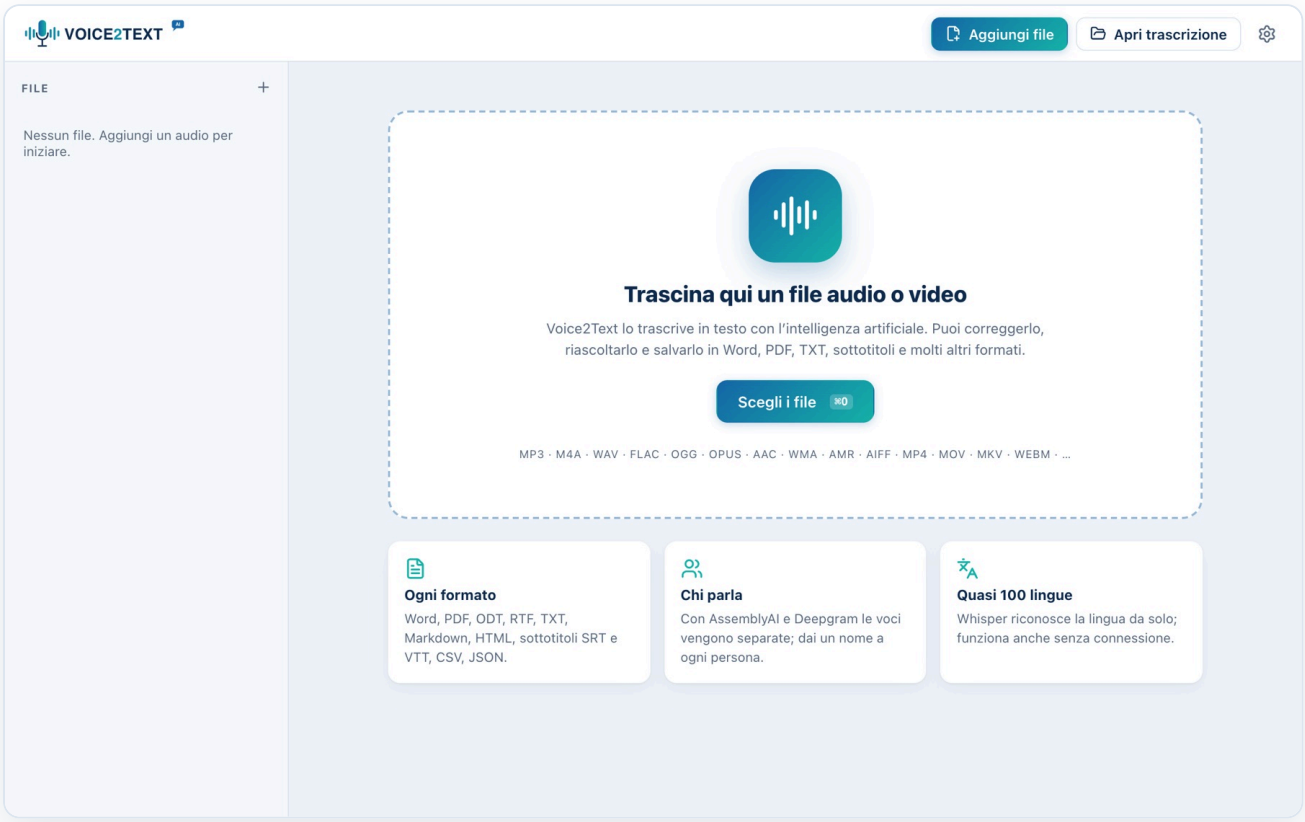


Figura 1. La schermata iniziale: trascina un file nel riquadro oppure usa «Scegli i file».

- **Barra in alto**: **Aggiungi file**, **Apri trascrizione** (un progetto `.v2t` salvato) e l'ingranaggio delle **Impostazioni**.
- **Elenco dei file** a sinistra: ogni file mostra lo stato (pronto, in coda, in trascrizione, trascritto, errore) e l'avanzamento. Il pallino dopo il nome indica modifiche non salvate.
- **Area centrale**: le impostazioni della trascrizione, l'avanzamento oppure il testo da correggere.

Aggiungere i file

Trascina uno o più file audio o video sulla finestra, oppure usa **Aggiungi file** (`⌘O` / `Ctrl+O`). Puoi aggiungerne quanti vuoi: vengono trascritti uno dopo l'altro. Trascinando un file `.v2t` riapri una trascrizione salvata.

04 Scegliere motore e lingua

> Prima di trascrivere scegli chi farà il lavoro: Whisper sul tuo computer oppure un servizio online.

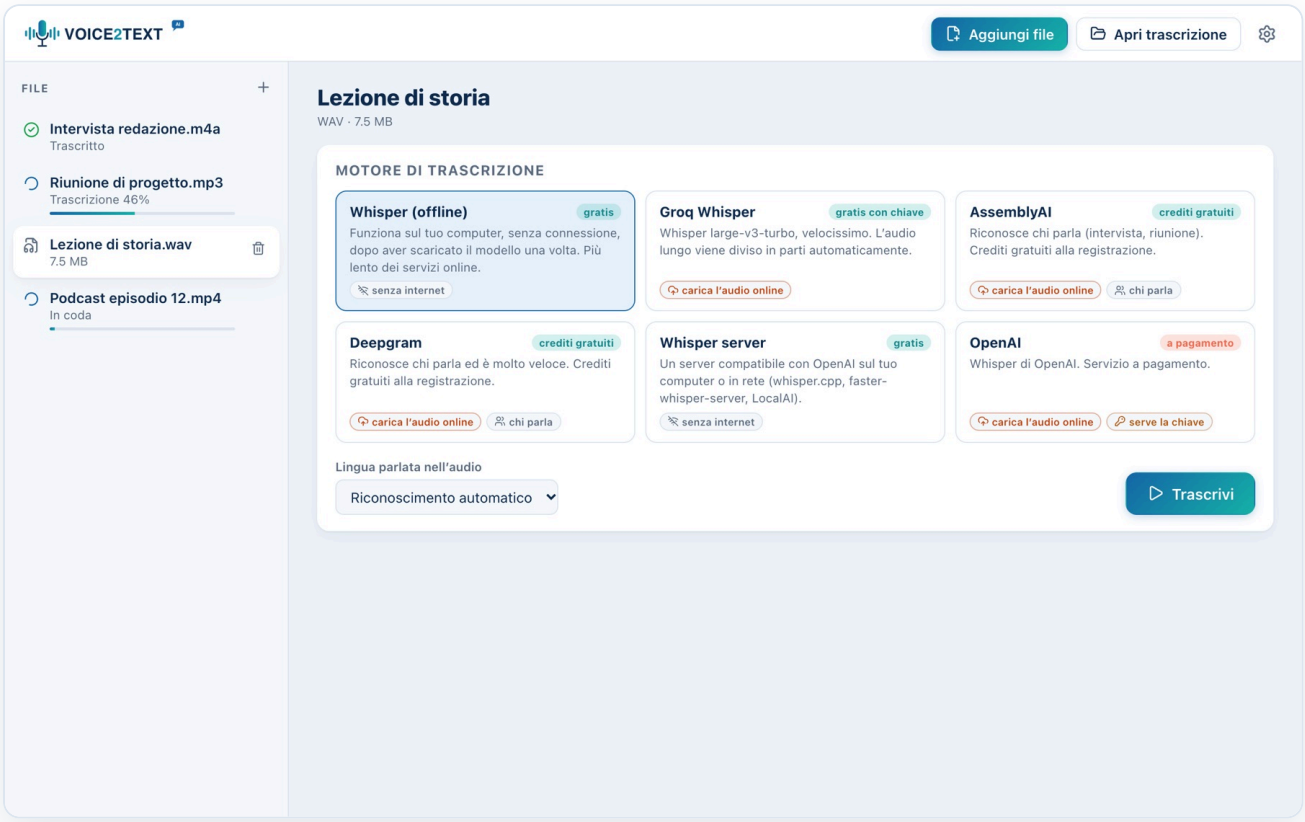


Figura 2. Scelta del motore e della lingua per il file selezionato.

Motore	Costo	Caratteristiche
Whisper (offline)	gratis	Funziona senza Internet; il modello si scarica una volta (da 45 a 600 MB). Più lento dei servizi online.
Groq Whisper	gratis con chiave	Velocissimo; l'audio lungo viene diviso in parti automaticamente.
AssemblyAI	crediti gratuiti	Riconosce chi parla. Ideale per interviste e riunioni.
Deepgram	crediti gratuiti	Riconosce chi parla ed è molto veloce.
Whisper server	gratis	Un server Whisper sul tuo computer o nella tua rete (whisper.cpp, faster-whisper-server, LocalAI).
OpenAI	a pagamento	Whisper di OpenAI.

La lingua dell'audio

Con **Riconoscimento automatico** Voice2Text capisce da solo la lingua parlata. Se la conosci, sceglila: il risultato è più preciso, soprattutto nei primi secondi.

Le chiavi dei servizi online

Groq, AssemblyAI, Deepgram e OpenAI chiedono una chiave personale: in **Impostazioni** → **Motori** trovi il campo per ognuno e il collegamento «ottieni la chiave». La registrazione a Groq è gratuita; AssemblyAI e Deepgram regalano crediti all'iscrizione.

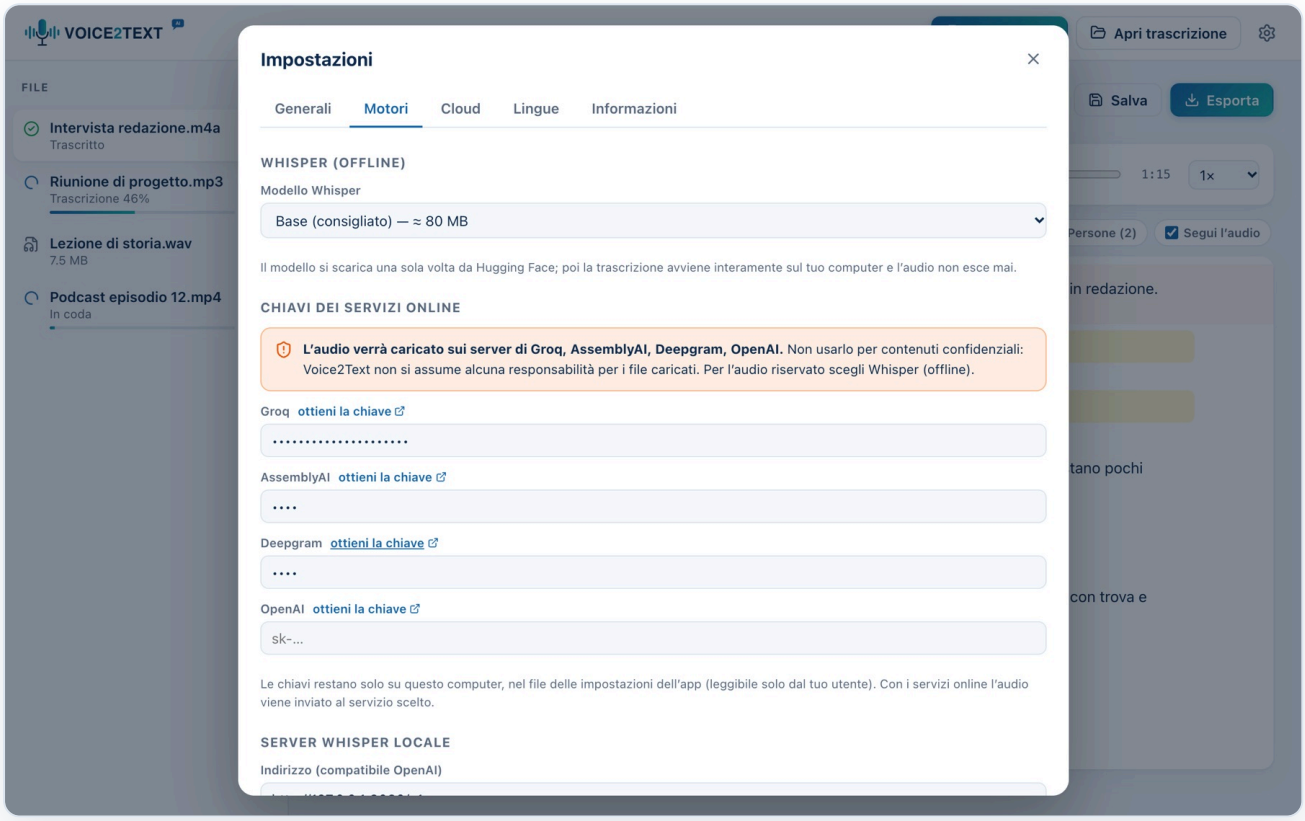


Figura 3. Impostazioni → Motori: modello Whisper e chiavi dei servizi.

Premi **Trascrivi** per iniziare. Con più file in attesa compare anche **Trascrivi tutti** nella barra laterale.

05 Servizi online: il tuo audio esce dal computer

> Quando usi un servizio di intelligenza artificiale online, il file audio viene caricato su Internet. Voice2Text te lo dice chiaramente e chiede la tua conferma.

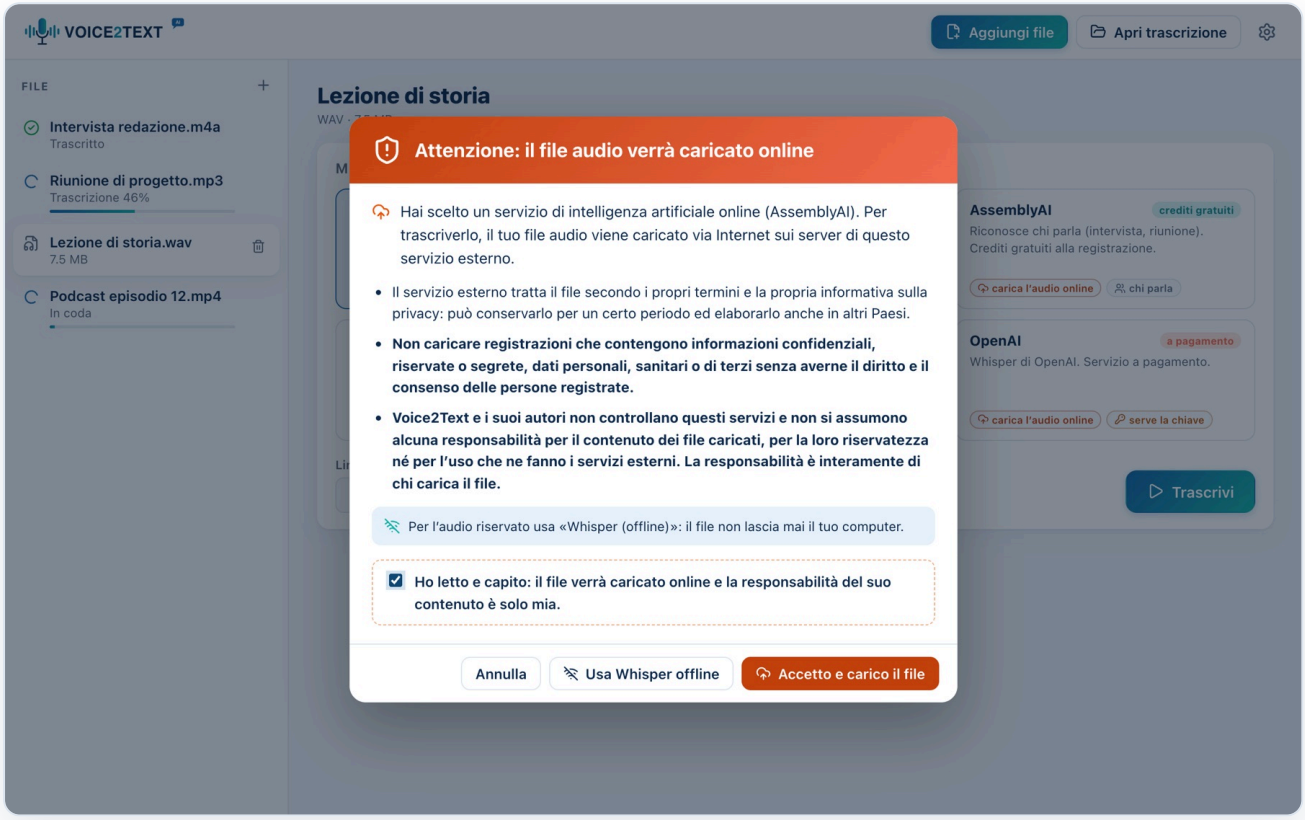


Figura 4. L'avviso obbligatorio prima del primo invio a un servizio online.

La prima volta che usi Groq, AssemblyAI, Deepgram o OpenAI compare questo avviso. Per proseguire devi spuntare la casella **«Ho letto e capito»**; in alternativa puoi passare subito a **Whisper offline**.

> **Il file audio viene caricato sui server del servizio esterno**, che lo tratta secondo i propri termini e la propria informativa sulla privacy: può conservarlo per un certo periodo ed elaborarlo anche in altri Paesi.
>
> **Non caricare registrazioni con informazioni confidenziali**, riservate o segrete, dati personali, sanitari o di terzi senza averne il diritto e il consenso delle persone registrate.
>
> **Voice2Text e i suoi autori non si assumono alcuna responsabilità** per il contenuto dei file caricati, per la loro riservatezza né per l'uso che ne fanno i servizi esterni. La responsabilità è interamente di chi carica il file.

L'accettazione vale per il singolo servizio e viene registrata con la data. Finché è scelto un servizio online, un riquadro arancione te lo ricorda. In **Impostazioni → Informazioni** puoi rileggere l'avviso e farlo ricomparire.

Audio riservato? Usa **Whisper (offline)**: la trascrizione avviene interamente sul tuo computer.

06 Durante la trascrizione

> Voice2Text mostra in ogni momento a che punto è il lavoro.

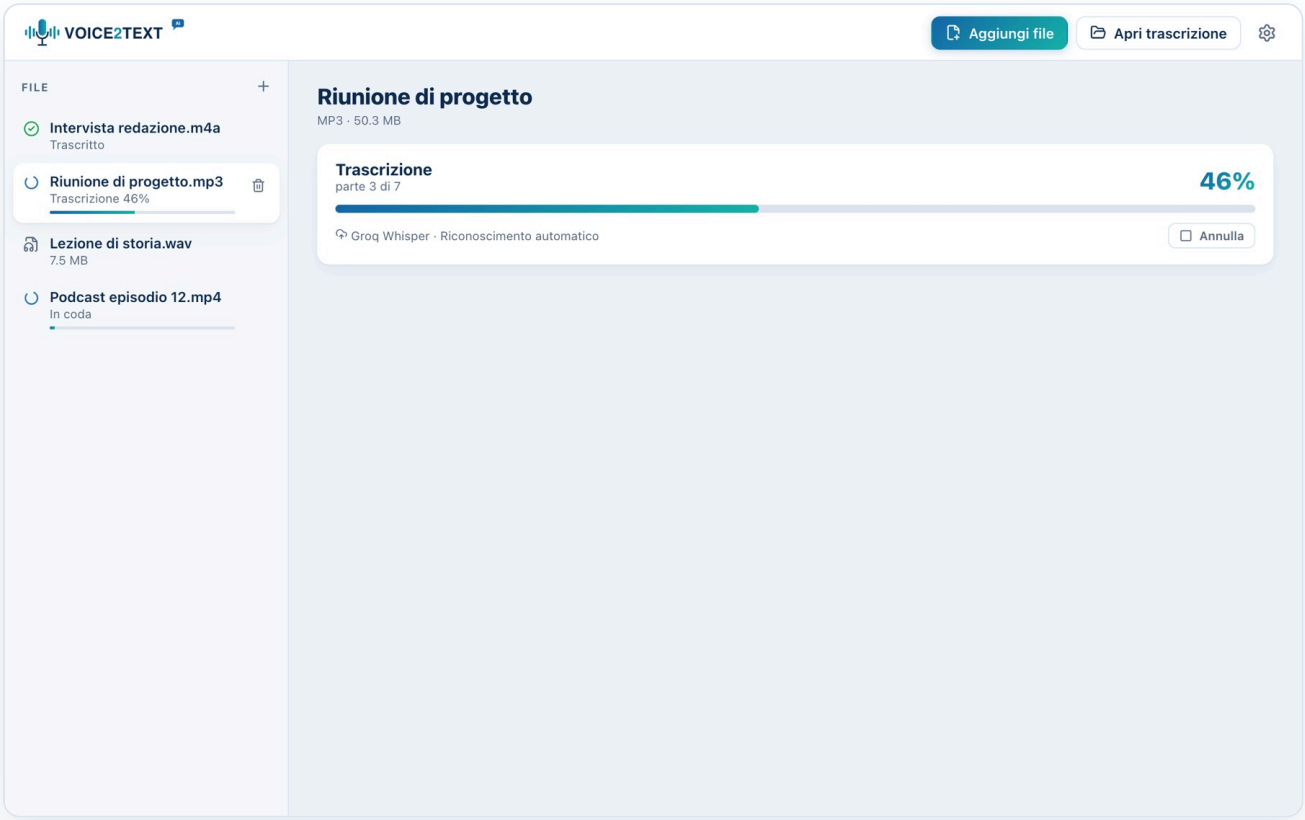


Figura 5. Una trascrizione in corso con Groq: avanzamento e parte in elaborazione.

- **Lettura dell'audio**: il file viene decodificato sul computer.
- **Preparazione del modello** (solo Whisper offline, la prima volta): il modello scelto viene scaricato e conservato.
- **Invio dell'audio** e **Trascrizione**: con i file lunghi vedi la parte in lavorazione («parte 3 di 7»).
- **In attesa del servizio**: con i piani gratuiti, se si supera il limite di richieste, Voice2Text aspetta e riprova da solo.

Puoi annullare in qualsiasi momento con **Annulla**, continuare a lavorare su altri file e chiudere la finestra della trascrizione: quando il testo è pronto ricevi una notifica.

07 Correggere la trascrizione

> Il testo è diviso in frasi, ognuna con il suo tempo e, se riconosciuta, la persona che parla.

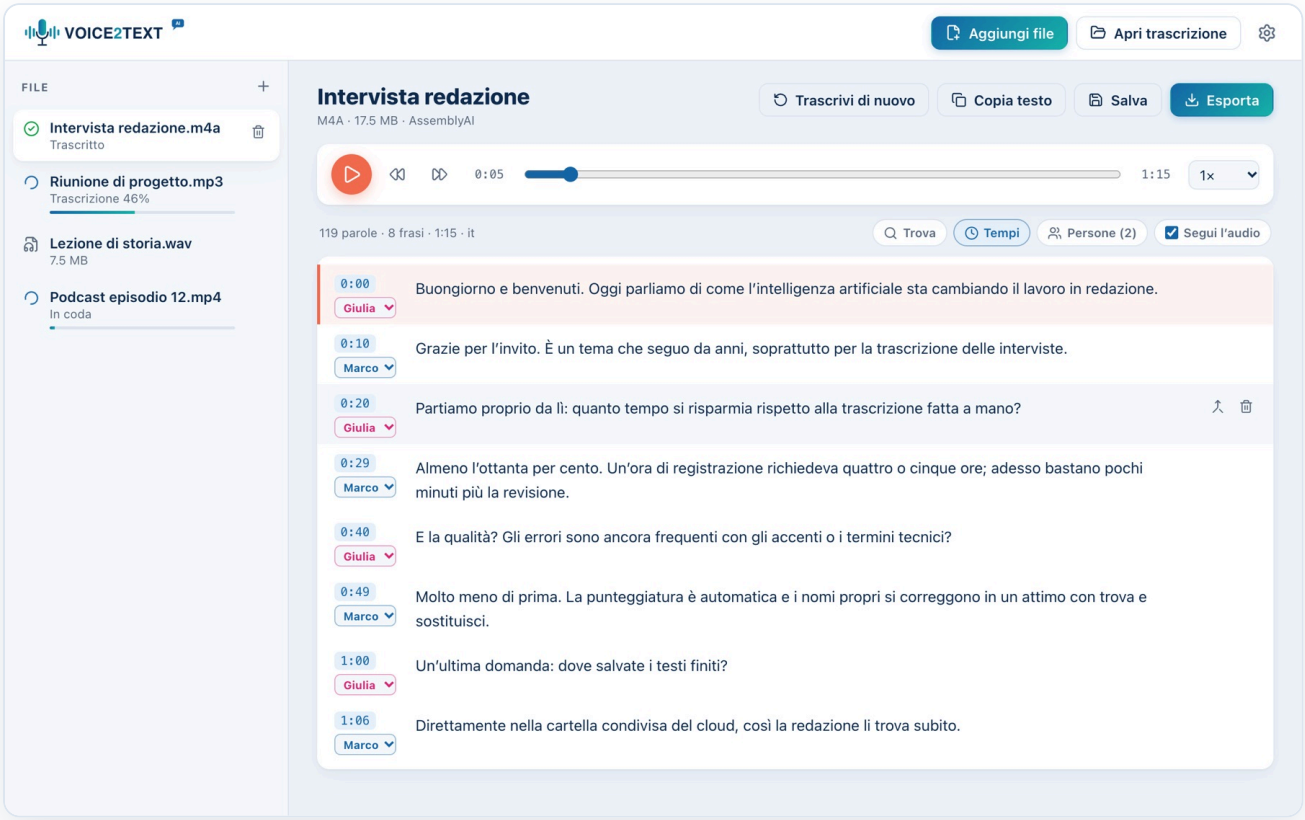


Figura 6. L'editor: lettore audio in alto e frasi modificabili; la frase in ascolto è evidenziata.

Ascoltare

Il lettore in alto riproduce l'audio originale: **Spazio** avvia e mette in pausa (quando non stai scrivendo), i pulsanti saltano di 5 secondi e puoi scegliere la velocità da 0,75× a 2×. Fai clic sul tempo di una frase per ascoltarla; con **Segui l'audio** la frase in riproduzione resta sempre visibile.

Modificare

- Fai clic su una frase e scrivi: le modifiche sono immediate.
- **Invio** divide la frase nel punto del cursore; **Maiusc+Invio** va a capo nella stessa frase.
- **Backspace** all'inizio di una frase la unisce alla precedente; il pulsante **Unisci** la unisce alla successiva.
- Il cestino elimina una frase (per esempio un rumore riconosciuto come parola).

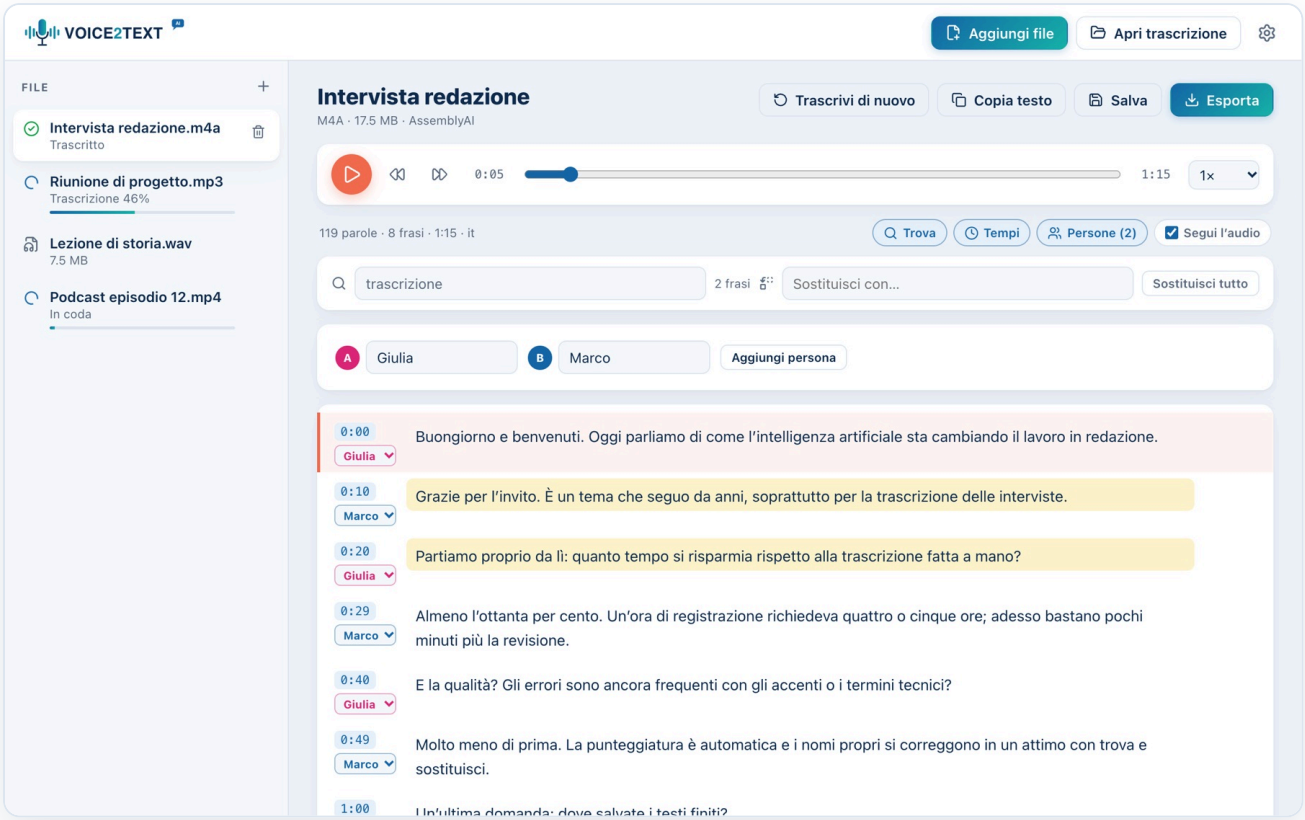


Figura 7. Trova e sostituisci, e i nomi delle persone che parlano.

Trova e sostituisci

Trova (`⌘F` / `Ctrl+F`) evidenzia le frasi che contengono la parola cercata; **Sostituisci tutto** corregge in un colpo un nome scritto male in tutta la trascrizione.

Persone

Con AssemblyAI e Deepgram ogni frase ha un'etichetta (Persona A, Persona B…). Nel pannello **Persone** scrivi il nome vero: compare ovunque, anche nei file esportati. Puoi cambiare la persona di una frase dal menu accanto al tempo, aggiungerne altre o assegnarle a mano anche quando il motore non le riconosce.

08 Salvare ed esportare

> Salva il progetto per riprendere le correzioni più tardi ed esporta il testo nel formato che ti serve.

Il progetto .v2t

Salva (⌘S / Ctrl+S) crea un file `.v2t` con la trascrizione, le correzioni e i nomi delle persone. Lo riapri con **Apri trascrizione** (⇧⌘O / Ctrl+Maiusc+O) o trascinandolo sulla finestra. ⇧⌘S salva con un altro nome.

Esportare

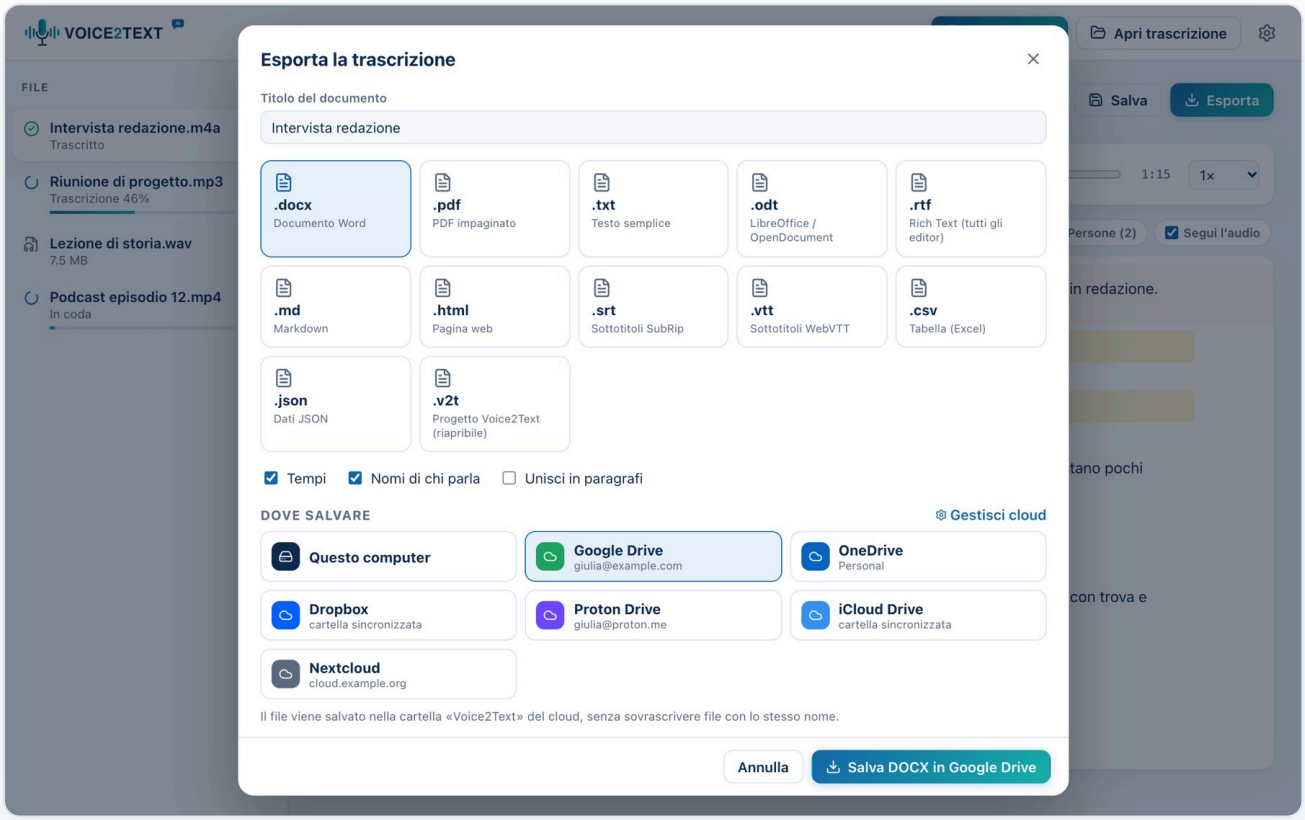


Figura 8. La finestra Esporta: formato, opzioni e destinazione.

Premi **Esporta** (⌘E / Ctrl+E), scrivi il titolo del documento e scegli il formato:

Formato	Ideale per
.docx Word · **.odt** LibreOffice · **.rtf**	documenti da modificare, impaginati con titolo e dati della registrazione
.pdf	documenti da stampare o inviare; si crea con la finestra di stampa del sistema («Salva come PDF»)
.txt · **.md** · **.html**	testo semplice, Markdown, pagina web
.srt · **.vtt**	sottotitoli per video (YouTube, Premiere, DaVinci, VLC…)
.csv · **.json**	tabelle (Excel, Numbers) e dati per altri programmi
.v2t	il progetto riapribile, anche da salvare nel cloud

Le opzioni scelgono se includere i **tempi**, i **nomi di chi parla** e se **unire le frasi in paragrafi**. Il testo arabo e le altre lingue da destra a sinistra vengono impaginati correttamente.

09 Salvare nel cloud

> Voice2Text salva i file direttamente nel tuo cloud usando le cartelle che le app dei servizi creano sul computer. Non devi collegare account né inserire password.

Come funziona

Quando installi l'app di Google Drive, OneDrive, Dropbox, Box, **Proton Drive**, iCloud Drive, pCloud, MEGA, Nextcloud o Sync.com, sul computer compare una cartella sincronizzata. Voice2Text la trova da solo: nella finestra **Esporta**, alla voce **Dove salvare**, scegli il cloud e premi il pulsante. Il file viene scritto nella sottocartella `Voice2Text` e l'app del servizio lo carica subito online.

- I file non vengono mai sovrascritti: se esiste già un file con lo stesso nome, il nuovo diventa «nome (2)».
- Puoi salvare nel cloud qualsiasi formato, compreso il progetto `.v2t`.
- Il PDF passa dalla finestra di stampa: lì scegli «Salva come PDF» e la cartella del cloud.

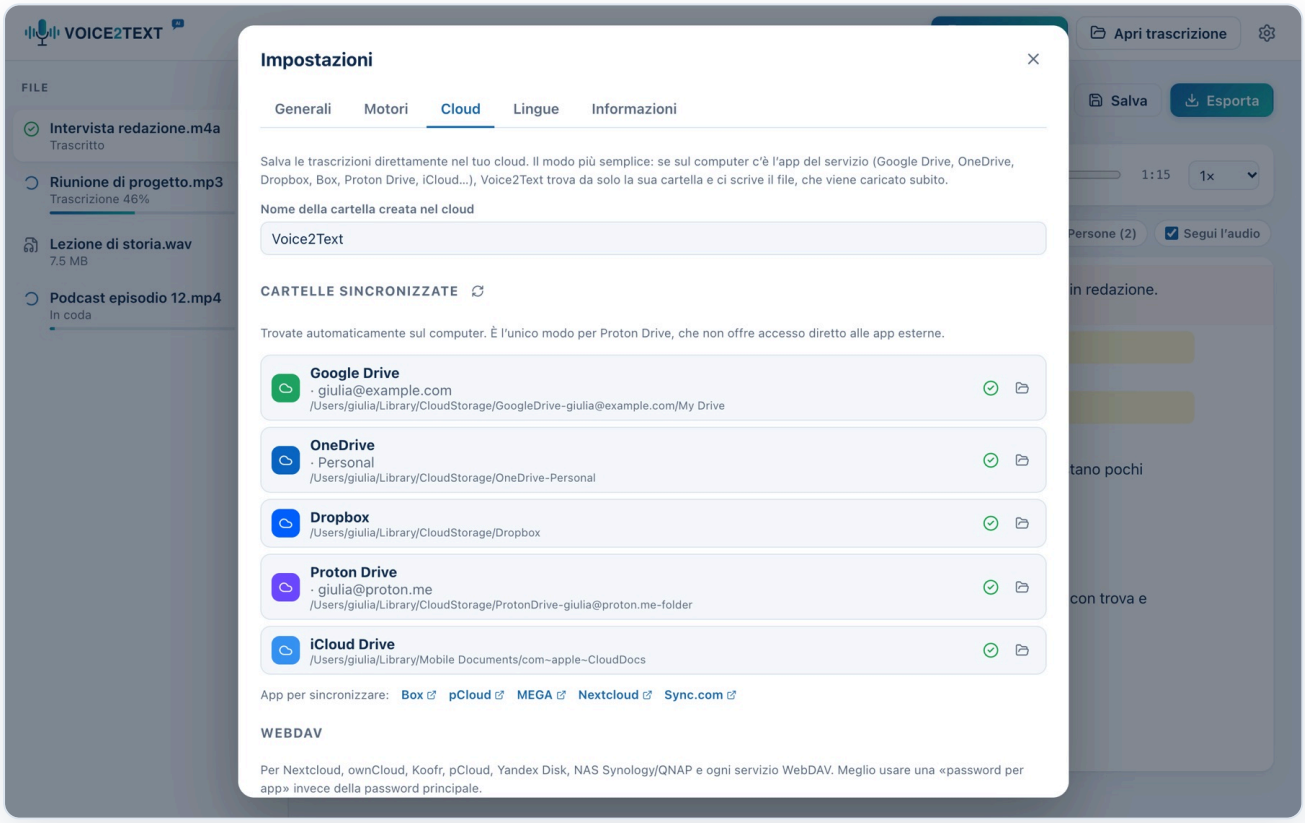


Figura 9. Impostazioni → Cloud: le cartelle trovate sul computer e il collegamento WebDAV.

Impostazioni → Cloud

- **Nome della cartella**: la sottocartella creata nel cloud (predefinita `Voice2Text`).
- **Cartelle sincronizzate**: l'elenco dei cloud trovati; l'icona della cartella la apre. Se manca un servizio, i collegamenti in basso portano alle app da installare.
- **Aggiungi un'altra cartella**: per altri cloud o per un disco di rete (NAS).
- **WebDAV**: per Nextcloud, ownCloud, Koofr, pCloud, Yandex Disk e NAS Synology/QNAP senza app installata. Scegli il servizio, inserisci indirizzo, utente e una password per app, poi premi **Prova la**

connessione.

> **Proton Drive** non permette ad altre app di accedere direttamente all’account: per questo Voice2Text usa la sua cartella sincronizzata, che offre la stessa cifratura end-to-end.

10 Impostazioni e lingue

L'ingranaggio in alto a destra (`⌘,` / `Ctrl+,`) apre le impostazioni.

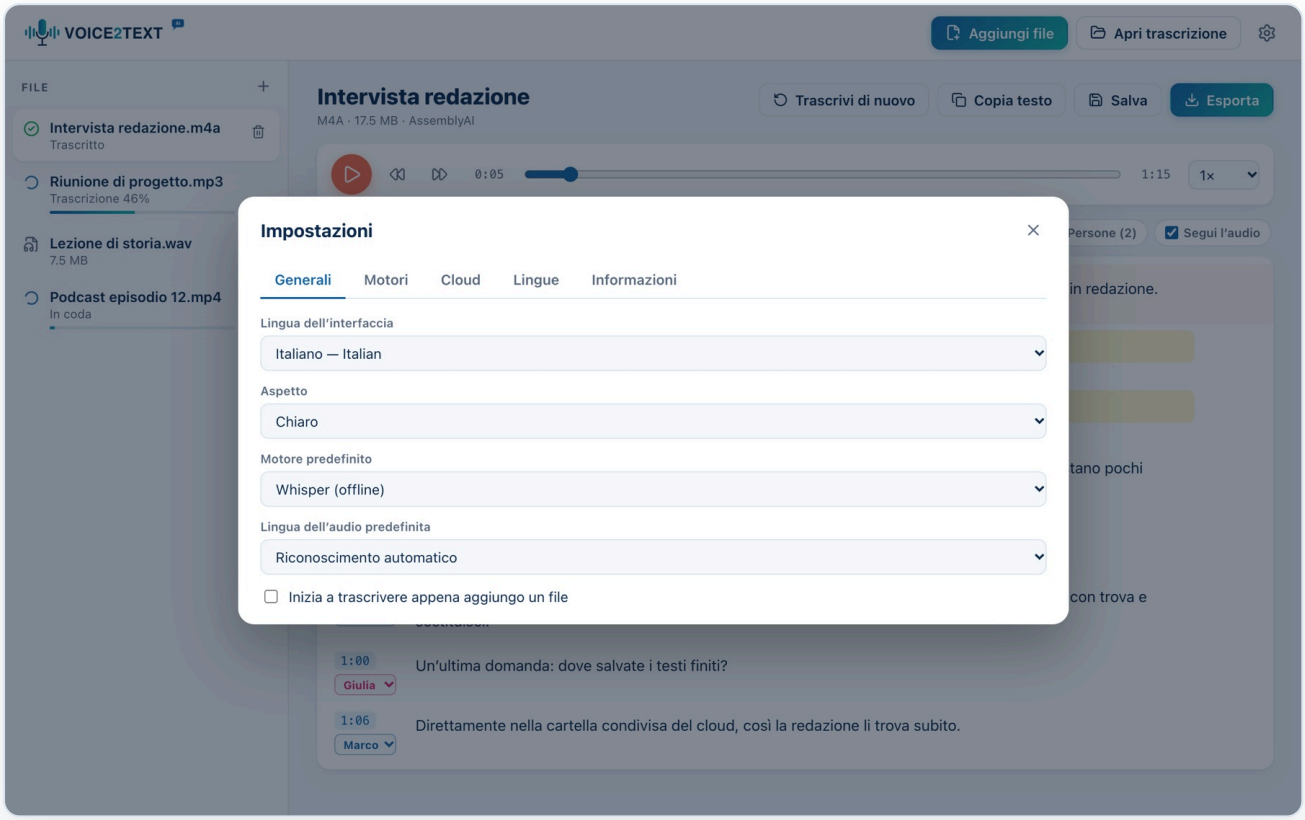


Figura 10. Impostazioni → Generali.

- **Generali**: lingua dell'interfaccia, aspetto chiaro, scuro o automatico, motore e lingua dell'audio predefiniti, avvio automatico della trascrizione.
- **Motori**: modello di Whisper offline (minimo, base, piccolo, grande turbo), spazio occupato dai modelli, chiavi dei servizi, server Whisper locale, presenza di ffmpeg.
- **Cloud**: cartelle sincronizzate e WebDAV (capitolo precedente).
- **Lingue**: i file di traduzione dell'interfaccia.
- **Informazioni**: versione, licenza, avviso sui servizi online.

Le lingue sono file XML

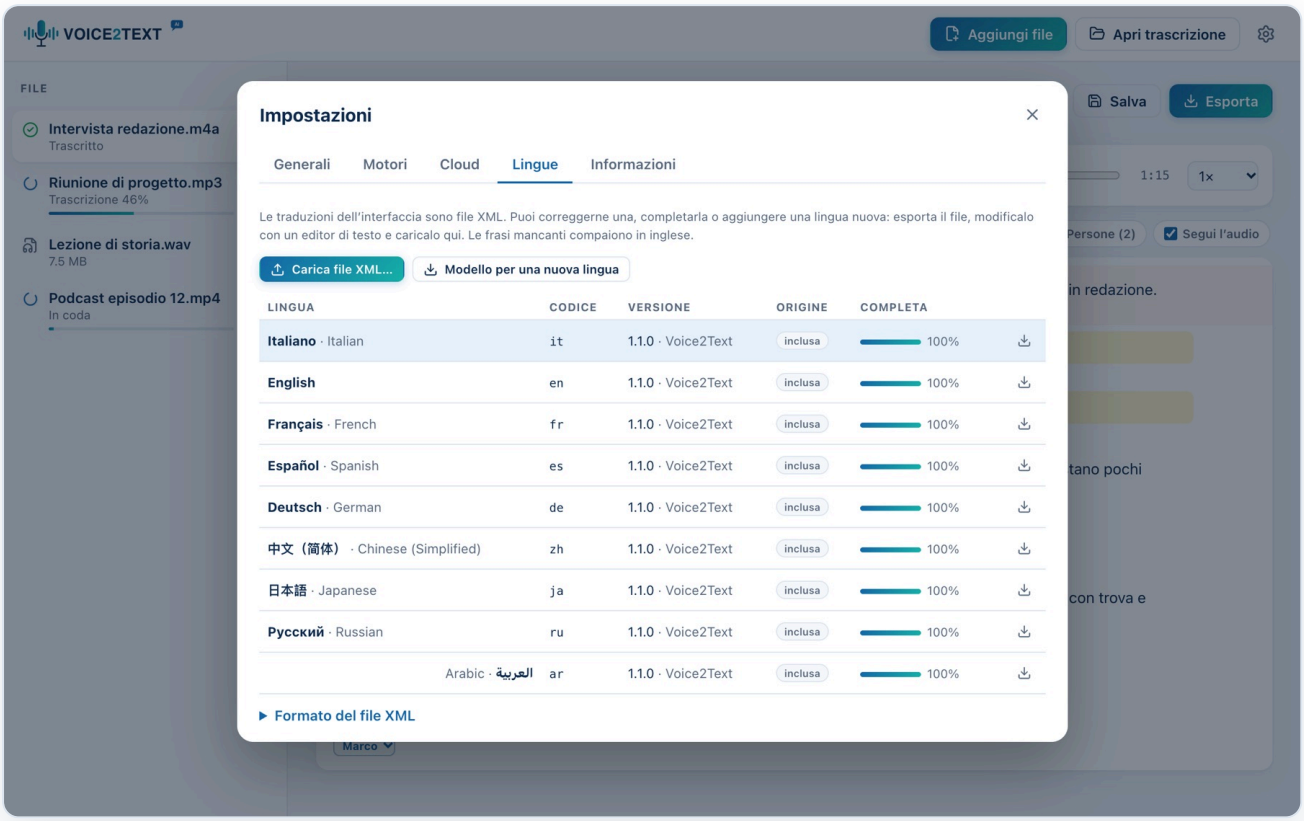


Figura 11. Impostazioni → Lingue: lingue disponibili, completezza e strumenti per i file XML.

Ogni lingua dell'interfaccia è un file XML. Per correggere una traduzione o aggiungerne una nuova:

1. Esporta il file della lingua (freccia nella riga) oppure scarica il **Modello per una nuova lingua**.
2. Aprilo con un editor di testo e traduci solo il testo tra i tag, lasciando uguali gli attributi `key` e i segnaposto tra graffe come `{n}`.
3. Premi **Carica file XML...**: la lingua compare subito nell'elenco e nel menu delle lingue.

```
<language code="pt" name="Português" english="Portuguese" dir="ltr" version="1.0">
  <string key="app.addFiles">Adicionar ficheiros</string>
  <string key="progress.piece">parte {i} de {n}</string>
</language>
```

Un file con lo stesso codice di una lingua inclusa la sostituisce (per le correzioni); puoi sempre tornare alla versione originale. Le frasi mancanti compaiono in inglese; la colonna **Completa** mostra quante sono tradotte. Per le lingue scritte da destra a sinistra usa `dir="rtl"`.

11 Scorciatoie da tastiera

> Su Windows e Linux usa **Ctrl** al posto di ⌘.

Tasti	Azione
⌘O	Aggiungi file audio o video
⇧⌘O	Apri una trascrizione salvata (.v2t)
⌘S · ⇧⌘S	Salva · Salva con nome
⌘E	Esporta
⌘F	Trova e sostituisci
⌘,	Impostazioni
Spazio	Riproduci / pausa (fuori dal testo)
Invio · **Maiusc+Invio**	Dividi la frase · A capo nella frase
Backspace a inizio frase	Unisci alla frase precedente
Esc	Esci dal testo · chiudi la finestra aperta

12 Domande frequenti

> Le risposte ai dubbi più comuni.

Il formato del mio file non viene letto.

Installa ffmpeg (capitolo Installazione) oppure usa AssemblyAI o Deepgram, che accettano qualsiasi formato perché ricevono il file originale.

Whisper offline è lento.

Scegli il modello **Base** o **Minimo** in Impostazioni → Motori. Il modello «Grande turbo» è il più preciso ma senza scheda grafica richiede tempo. Per file lunghi Groq è molto più veloce.

Il testo è in un'altra lingua o tradotto.

Indica la lingua dell'audio invece del riconoscimento automatico e usa **Trascrivi di nuovo**.

«Chiave del servizio non valida».

Controlla la chiave in Impostazioni → Motori: copiala di nuovo dal sito del servizio, senza spazi.

Il mio cloud non compare.

Verifica che l'app del servizio sia installata e sincronizzata, poi premi **Cerca di nuovo** in Impostazioni → Cloud. In alternativa aggiungi la cartella a mano.

Il lettore dice «audio non disponibile».

Il file audio originale è stato spostato o rinominato dopo il salvataggio del progetto. Il testo resta disponibile; per riascoltare, rimetti il file al suo posto.

Dove sono le mie chiavi e le mie impostazioni?

Solo sul tuo computer, nel file delle impostazioni dell'app, leggibile soltanto dal tuo utente. Non vengono mai inviate a Voice2Text.



VOICE2TEXT

Dalla voce al testo, in ogni lingua.

Voice2Text è un programma di overminder. Licenza PolyForm Strict 1.0.0: uso non commerciale consentito; ridistribuzione, modifiche e opere derivate non consentite.

© 2026 overminder · Versione 1.1