



SCHEDA TECNICA · ACQUISIRE · VOICE STT

Trascrivere le riunioni parola per parola? D'ora in poi ogni registrazione torna come testo.

Voice STT trasforma le registrazioni di riunioni, chiamate e dettature in testo. Riconosce da solo la lingua, e il livello Ultra separa i parlanti. Pensata per team di ufficio, vendite e progetti, nell'app e tramite API.

IN SINTESI

File audio e video fino a 50 MB, riconoscimento automatico della lingua, tre livelli di qualità, separazione dei parlanti in Ultra e timestamp delle parole, SRT e VTT tramite API.

01 · DI COSA SI TRATTA

Prima la riunione, poi la battitura? Ora il testo arriva da solo.

Il verbale della riunione di progetto deve partire oggi, e la registrazione dura un'ora. Con Voice STT carica il file e riceve il testo indietro, separato per partecipante in Ultra, pronto da copiare o tradurre.

INGRESSO

Audio (MP3, WAV, OGG, FLAC, M4A) e video (MP4, WebM, AVI, MOV)
fino a 50 MB

LINGUE

Riconosciute automaticamente; 100 codici lingua via API, 15 lingue da scegliere nell'app

USCITA

Trascrizione in testo semplice, in Ultra come dialogo per parlante;
download come TXT

ACCESSO

App sotto

Media AI › Studio Vocale
, REST API e MCP

**Lingua riconosciuta automaticamente**

Voice STT riconosce da solo la lingua della registrazione. Se la conosce, può impostarla: tramite API da 100 codici lingua.

Tre livelli di qualità

Basic restituisce velocemente testo semplice, Premium aggiunge segmenti con timestamp delle parole, Ultra separa anche i parlanti. Il livello si trova sotto **Qualità di trascrizione**.

Separare i parlanti

In Ultra, Voice STT riconosce i partecipanti e mostra la trascrizione come dialogo. I parlanti si rinominano con un clic.

Vocabolario

Con **Lessico tecnico**, Voice STT riconosce con maggiore precisione i termini tecnici delle Sue basi di conoscenza e dei listini. Tramite API aggiunge le proprie hotword dal livello Premium.

Sottotitoli tramite API

Tramite API, dal livello Premium, restituisce SRT e VTT più i timestamp delle parole, e Ultra aggiunge RTTM per l'analisi dei parlanti.

Passaggio a Translate Studio

Usa in **Translate Studio** passa la trascrizione direttamente allo Studio di traduzione, senza copiare e incollare.

02 · COME FUNZIONA

Trascini il file, scelga il livello. Usi la trascrizione.

In Voice Studio bastano quattro passaggi. Tramite REST API e MCP invia i file audio direttamente dai Suoi processi.

1

Apra Voice Studio

Sotto **Media AI > Studio Vocale** scelga la scheda **Da parlato a testo (STT)**. Dalla pagina iniziale si arriva qui anche con **Strumenti rapidi > Media > Riconoscimento vocale**.

2

Trascini il file

Trascini il file audio o video nel campo. Lasci la lingua su Auto oppure la scelga.

3

Scelga livello e opzioni

Sotto **Qualità di trascrizione** scelga Basic, Premium o Ultra. In Ultra può impostare il numero dei parlanti sotto **Riconoscimento degli interlocutori**.

4

Usi la trascrizione

La copi, la salvi come file di testo con **Scarica come TXT** oppure la passa allo Studio di traduzione con un clic.

DATI TECNICI

CARATTERISTICA	DETTAGLI
Formati di ingresso	MP3, WAV, OGG, FLAC, M4A, MP4, WebM, AVI, MOV; nell'app fino a 50 MB, un file per avvio
Lingue	riconoscimento automatico; 100 codici lingua via API, 15 scelte fisse nell'app
Livelli di qualità	Basic (predefinito API), Premium, Ultra (predefinito app)
Uscita via API	output_format json, srt e vtt da Premium, rttm in Ultra
Separazione dei parlanti	in Ultra, min_speakers e max_speakers facoltativi da 1 a 20
REST API	POST /job/add/paperoffice_voice___stt con audio_file
MCP	po_stt_transcribe su /mcp-full

03 · IN PRATICA

Riunione, dettatura, video formativo. Ognuno come testo pronto.

Verbale di una riunione di progetto

Il responsabile di progetto carica la registrazione in Ultra, rinomina Speaker 1 e 2 con i loro nomi e copia il dialogo nel verbale.

Dettatura dal campo

Un perito detta al telefono durante il tragitto e carica il file M4A in ufficio. Basic restituisce il testo semplice per il rapporto, e il vocabolario conosce i nomi dei prodotti.

Sottotitoli per un video formativo

Lo sviluppo invia il video tramite API con `quality=premium` e `output_format=srt` e riceve indietro il file dei sottotitoli con i timestamp.

DOMANDE FREQUENTI

Quali lingue riconosce Voice STT?

La lingua viene riconosciuta automaticamente. Tramite API può impostarla tra 100 codici lingua, mentre l'app offre 15 lingue come scelte fisse.

Voice STT distingue i parlanti?

Sì, in Ultra. La trascrizione appare come dialogo e i parlanti possono essere rinominati. Tramite API ottiene anche il numero dei parlanti, le sovrapposizioni e RTTM.

Posso ottenere i sottotitoli come SRT o VTT?

Sì, tramite API dal livello Premium con `output_format=srt` o `vtt`. Nell'app salva la trascrizione come file TXT.

Come viene fatturato?

In base all'utilizzo, a seconda del livello di qualità. I prezzi attuali sono nella pagina dei prezzi.

INIZIA ORA

Faccia trascrivere per Lei la prossima registrazione.

Acceda al Suo account PaperOffice, apra **Media AI > Studio Vocale** e scelga **Da parlato a testo (STT)**. Trascini il file, scelga un livello, avvii la trascrizione.
paperoffice.ai/it/ai-voice-stt-riconoscimento-vocale-trascrizione-stt



MAGGIORI INFORMAZIONI SUL MODULO

paperoffice.ai/it/ai-voice-stt-riconoscimento-vocale-trascrizione-stt

PREZZI

paperoffice.ai/en/pricing

CONTATTI

paperoffice.ai/it/contatto