



TEHNIČKI LIST • SNIMANJE • GLAS STT

Sastanke kucate reč po reč? Od sada se svaka snimka vraća kao tekst.

Glas STT pretvara snimke sastanaka, poziva i diktata u tekst. Jezik sam prepoznaje, a nivo Ultra razdvaja govornike. Napravljen za kancelarijske, prodajne i projektne timove, u aplikaciji i preko API-ja.

NA PRVI POGLED

Audio i video datoteke do 50 MB, automatsko prepoznavanje jezika, tri nivoa kvaliteta, razdvajanje govornika u Ultra nivou i vremenske oznake reči, SRT i VTT preko API-ja.

01 · O ČEMU SE RADI

Prvo sastanak, pa kucanje? Sada tekst stiže sam.

Zapisnik sa projektnog sastanka mora da izađe danas, a snimak traje sat vremena. Uz Glas STT otpremite datoteku i dobijate tekst nazad, odvojen po učesniku u Ultra nivou, spreman za kopiranje ili prevod.

UNOS

Audio (MP3, WAV, OGG, FLAC, M4A) i video (MP4, WebM, AVI, MOV) do 50 MB

JEZICI

Prepoznaju se automatski; 100 kodova jezika preko API-ja, 15 jezika za izbor u aplikaciji

IZLAZ

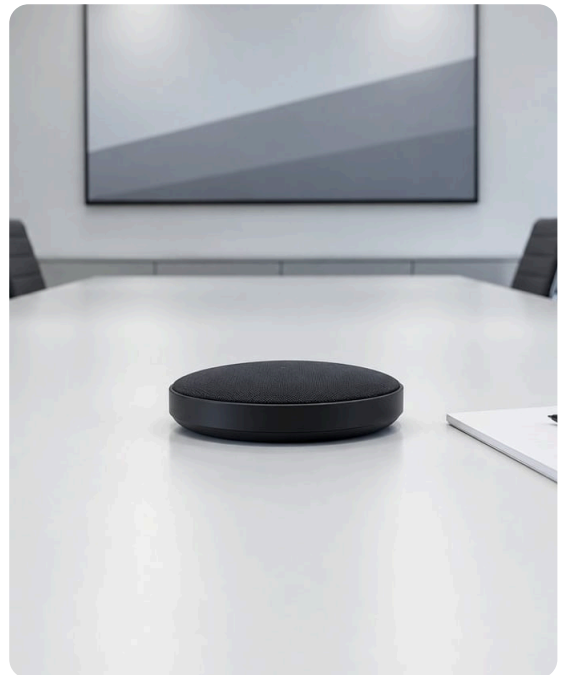
Transkript kao običan tekst, u Ultra nivou kao dijalog po govorniku; preuzimanje kao TXT

PRISTUP

Aplikacija pod

Media AI › Glasovni studio

, REST API i MCP

**Jezik se prepoznaje sam**

Glas STT sam prepoznaje jezik snimka. Ako ga znate, možete ga podesiti preko API-ja, iz 100 kodova jezika.

Tri nivoa kvaliteta

Basic vraća brz običan tekst, Premium dodaje segmente sa vremenskim oznakama reči, a Ultra razdvaja i govornike. Nivo se nalazi pod **Kvalitet transkripcije**.

Razdvajanje govornika

U Ultra nivou Glas STT prepoznaje učesnike i prikazuje transkript kao dijalog. Govornike preimenujete jednim klikom.

Rečnik

Uz **Stručni rečnik**, Glas STT tačnije prepoznaje stručne pojmove iz vaših baza znanja i cenovnika. Preko API-ja dodajete sopstvene ključne reči od nivoa Premium.

Titlovi preko API-ja

Preko API-ja Premium i viši nivoi vraćaju SRT i VTT i vremenske oznake reči, a Ultra dodaje RTTM za analizu govornika.

Dalje u Translate Studio

Koristite u **Translate Studio** prebacuje transkript direktno u Translate Studio, bez kopiranja i lepljenja.

02 · KAKO FUNKCIONIŠE

Prevucite datoteku, izaberite nivo. Iskoristite transkript.

U Voice Studiju to su četiri koraka. Preko REST API-ja i MCP-a audio datoteke šaljete direktno iz sopstvenih procesa.

1

Otvorite Voice Studio

Pod **Media AI** » **Glasovni studio** izaberite karticu **Govor u tekst (STT)**. Sa početne stranice vas i **Brzi alati** » **Media** » **Govor u tekst** vodi tamo.

2

Prevucite datoteku

Prevucite audio ili video datoteku u polje. Ostavite jezik na Auto ili ga izaberite.

3

Izaberite nivo i opcije

Pod **Kvalitet transkripcije** izaberite Basic, Premium ili Ultra. U Ultra nivou broj govornika možete podesiti pod **Prepoznavanje sagovornika**.

4

Iskoristite transkript

Kopirajte ga, sačuvajte kao tekstualnu datoteku pomoću **Preuzmi kao TXT** ili ga jednim klikom prosledite u Translate Studio.

TEHNIČKI PODACI

FUNKCIJA	DETALJI
Ulazni formati	MP3, WAV, OGG, FLAC, M4A, MP4, WebM, AVI, MOV; u aplikaciji do 50 MB, jedna datoteka po pokretanju
Jezici	automatsko prepoznavanje; 100 kodova jezika preko API-ja, 15 fiksnih izbora u aplikaciji
Nivoi kvaliteta	Basic (podrazumevano za API), Premium, Ultra (podrazumevano u aplikaciji)
Izlaz preko API-ja	output_format json, srt i vtt od nivoa Premium, rttm u Ultra
Razdvajanje govornika	u Ultra nivou, opciono min_speakers i max_speakers od 1 do 20
REST API	POST /job/add/paperoffice_voice___stt sa audio_file
MCP	po_stt_transcribe na /mcp-full

03 · U PRAKSI

Sastanak, diktat, video za obuku. Svaki kao gotov tekst.

Zapisnik sa projektnog sastanka

Voda projekta otprema snimak u Ultra nivou, preimenuje Govornik 1 i Govornik 2 u njihova imena i prekopira dijalog u zapisnik.

Diktat sa terena

Geodetičar diktira telefonom u pokretu i u kancelariji otprema M4A datoteku. Basic vraća običan tekst za izveštaj, a rečnik poznaje nazive proizvoda.

Titlovi za video za obuku

Razvoj šalje video preko API-ja sa quality=premium i output_format=srt i dobija nazad datoteku titlova sa vremenskim oznakama.

ČESTO POSTAVLJANA PITANJA

Koje jezike Glas STT prepoznaje?

Jezik se prepoznaje automatski. Preko API-ja ga možete podesiti iz 100 kodova jezika, a aplikacija nudi 15 jezika kao fiksne izbore.

Može li Glas STT da razlikuje govornike?

Da, u Ultra nivou. Transkript se prikazuje kao dijalog, a govornike možete preimenovati. Preko API-ja dobijate i broj govornika, preklapanja i RTTM.

Mogu li da dobijem titlove kao SRT ili VTT?

Da, preko API-ja od nivoa Premium sa output_format=srt ili vtt. U aplikaciji transkript čuvate kao TXT datoteku.

Kako se naplaćuje?

Prema potrošnji, zavisno od nivoa kvaliteta. Aktuelne cene su na stranici sa cenama.

POČNITE

Neka vaš sledeći snimak bude pretvoren u tekst umesto vas.

Prijavite se na svoj PaperOffice nalog, otvorite [Media AI > Glasovni studio](#) i izaberite [Govor u tekst \(STT\)](#). Prevucite datoteku, izaberite nivo i pokrenite transkripciju.

paperoffice.ai/sr/ai-voice-stt-govor-u-tekst-transkripcija-stt



VIŠE O MODULU

paperoffice.ai/sr/ai-voice-stt-govor-u-tekst-transkripcija-stt

CENE

paperoffice.ai/en/pricing

KONTAKT

paperoffice.ai/sr/kontakt