



DATABLAD · FÅNGA UPP · VOICE STT

Skriver du av möten ord för ord? Från nu blir varje inspelning text.

Voice STT gör inspelningar av möten, samtal och diktat till text. Språket identifieras automatiskt, och nivån Ultra skiljer talarna åt. Utformad för kontors-, försäljnings- och projektteam, i appen och via API.

PÅ EN BLICK

Ljud- och videofiler upp till 50 MB, automatisk språkidentifiering, tre kvalitetsnivåer, talarseparering i Ultra samt ordtidsstämplar, SRT och VTT via API.

01 · VAD DET HANDLAR OM

Först mötet, sedan skrivandet? Nu kommer texten av sig själv.

Protokollet från projektmötet ska iväg i dag, och inspelningen är en timme lång. Med Voice STT laddar du upp filen och får texten tillbaka, uppdelad per deltagare i Ultra, redo att kopieras eller översättas.

INDATA

Ljud (MP3, WAV, OGG, FLAC, M4A) och video (MP4, WebM, AVI, MOV) upp till 50 MB

SPRÅK

Identifieras automatiskt; 100 språkkoder via API, 15 språk att välja i appen

RESULTAT

Avskrift som ren text, i Ultra som dialog per talare; nedladdning som TXT

ÅTKOMST

App under

Media AI › Röststudio

, REST-API och MCP

**Språk identifieras automatiskt**

Voice STT känner själv igen inspelningens språk. Om du vet det kan du ange det, via API från 100 språkkoder.

Tre kvalitetsnivåer

Basic ger snabb ren text, Premium lägger till segment med ordtidsstämplar, Ultra skiljer dessutom talarna åt. Nivån finns under **Transkriptionskvalitet**.

Skilj talarna åt

I Ultra identifierar Voice STT deltagarna och visar avskriften som en dialog. Byt namn på talarna med ett klick.

Ordförråd

Med **Fackordförråd** känner Voice STT igen tekniska termer från dina kunskapsbaser och prislistor bättre. Via API lägger du till egna nyckelord från Premium.

Undertexter via API

Via API ger Premium och uppåt SRT och VTT samt ordtidsstämplar, och Ultra lägger till RTTM för talaranalys.

Vidare till Translate Studio

Använd i **Translate Studio** skickar avskriften direkt till Translate Studio, utan att du behöver kopiera och klistra.

02 · SÅ FUNGERAR DET

Släpp filen, välj nivå. Använd avskriften.

I Voice Studio tar det fyra steg. Via REST-API och MCP skickar du ljudfiler direkt från dina egna processer.

1 Öppna Voice Studio

Under **Media AI** > **Röststudio** väljer du fliken **Tal till text (STT)**. Från startsidan tar **Snabbverktyg** > **Media** > **Tal till text** dig dit också.

2 Släpp filen

Dra ljud- eller videofilen till fältet. Låt språket stå på Auto eller välj det.

3 Välj nivå och inställningar

Under **Transkriptionskvalitet** väljer du Basic, Premium eller Ultra. I Ultra kan du ange antalet talare under **Identifiering av samtalspartners**.

4 Använd avskriften

Kopiera den, spara den som textfil med **Ladda ner som TXT** eller skicka den till Translate Studio med ett klick.

TEKNISKA DATA

FUNKTION	DETALJER
Indataformat	MP3, WAV, OGG, FLAC, M4A, MP4, WebM, AVI, MOV; i appen upp till 50 MB, en fil per körning
Språk	automatisk identifiering; 100 språkkoder via API, 15 fasta val i appen
Kvalitetsnivåer	Basic (API-standard), Premium, Ultra (appens standard)
Utdata via API	output_format json, srt och vtt från Premium, rttm i Ultra
Talarseparering	i Ultra, valfria min_speakers och max_speakers från 1 till 20
REST-API	POST /job/add/paperoffice_voice___stt med audio_file
MCP	po_stt_transcribe på /mcp-full

03 · I PRAKTIKEN

Möte, diktat, utbildningsfilm. Var och en som färdig text.

Protokoll från ett projektmöte

Projektledaren laddar upp inspelningen i Ultra, byter namn på Talare 1 och 2 till deras namn och kopierar dialogen in i protokollet.

Diktat från fältet

En besiktningsman diktar in i telefonen på vägen och laddar upp M4A-filen på kontoret. Basic ger ren text till rapporten, och ordförandet känner igen produktnamnen.

Undertexter till en utbildningsfilm

Utvecklingen skickar videon via API med `quality=premium` och `output_format=srt` och får tillbaka undertextfilen med tidsstämplar.

VANLIGA FRÅGOR

Vilka språk känner Voice STT igen?

Språket identifieras automatiskt. Via API kan du ange det med 100 språkkoder, och appen erbjuder 15 språk som fasta val.

Kan Voice STT skilja talare åt?

Ja, i Ultra. Avskriften visas som en dialog, och talare kan byta namn. Via API får du dessutom antalet talare, överlappningar och RTTM.

Kan jag få undertexter som SRT eller VTT?

Ja, via API från Premium-nivån med `output_format=srt` eller `vtt`. I appen sparar du avskriften som TXT-fil.

Hur faktureras det?

Efter användning, beroende på kvalitetsnivån. Aktuella priser finns på prissidan.

KOM IGÅNG

Låt nästa inspelning skrivas ut åt dig.

Logga in på ditt PaperOffice-konto, öppna **Media AI** > **Röststudio** och välj **Tal till text (STT)**. Släpp filen, välj nivå och starta transkriberingen.

paperoffice.ai/sv/ai-voice-stt-tal-till-text-transkription-stt



MER OM MODULEN

paperoffice.ai/sv/ai-voice-stt-tal-till-text-transkription-stt

PRISER

paperoffice.ai/en/pricing

KONTAKT

paperoffice.ai/sv/kontakt