



FICHE TECHNIQUE · CAPTURER · VOICE STT

Retranscrire les réunions mot à mot ? Désormais, chaque enregistrement revient sous forme de texte.

Voice STT transforme les enregistrements de réunions, d'appels et de dictées en texte. Il détecte seul la langue, et le niveau Ultra sépare les locuteurs. Conçu pour les équipes bureau, commerciales et projet, dans l'application et via API.

EN BREF

Fichiers audio et vidéo jusqu'à 50 Mo, détection automatique de la langue, trois niveaux de qualité, séparation des locuteurs en Ultra, ainsi que les horodatages des mots, SRT et VTT via API.

01 · DE QUOI IL S'AGIT

D'abord la réunion, puis la saisie ? Désormais, le texte arrive tout seul.

Le compte rendu de la réunion de projet doit partir aujourd'hui, et l'enregistrement dure une heure. Avec Voice STT, vous téléversez le fichier et récupérez le texte, séparé par intervenant en Ultra, prêt à être copié ou traduit.

ENTRÉE

Audio (MP3, WAV, OGG, FLAC, M4A) et vidéo (MP4, WebM, AVI, MOV)
jusqu'à 50 Mo

LANGUES

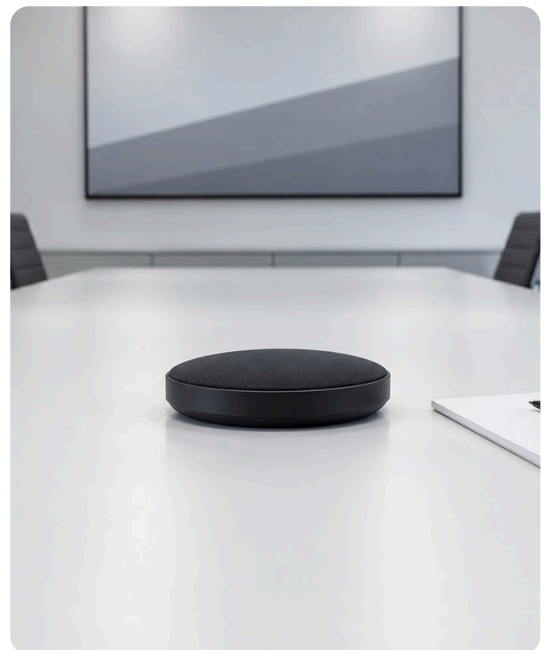
Détection automatique ; 100 codes de langue via API, 15 langues au choix dans l'application

SORTIE

Transcription en texte brut, en Ultra sous forme de dialogue par locuteur ; téléchargement en TXT

ACCÈS

Application sous
Media AI › Studio vocal
, API REST et MCP

**Langue détectée automatiquement**

Voice STT détecte seul la langue de l'enregistrement. Si vous la connaissez, vous pouvez la définir, via API parmi 100 codes de langue.

Trois niveaux de qualité

Basic renvoie rapidement un texte brut, Premium ajoute des segments avec horodatage des mots, Ultra sépare aussi les locuteurs. Le niveau se règle sous **Qualité de transcription**.

Séparer les locuteurs

En Ultra, Voice STT détecte les participants et affiche la transcription sous forme de dialogue. Renommez les locuteurs en un clic.

Vocabulaire

Avec le **Vocabulaire technique**, Voice STT reconnaît avec plus de précision les termes techniques de vos bases de connaissances et listes de prix. Via API, vous ajoutez vos propres mots-clés prioritaires à partir du niveau Premium.

Sous-titres via API

Via API, à partir du niveau Premium, le service renvoie SRT et VTT ainsi que les horodatages des mots, et Ultra ajoute RTTM pour l'analyse des locuteurs.

Vers Translate Studio

Utiliser dans **Translate Studio** transmet la transcription directement à Translate Studio, sans copier-coller.

02 · FONCTIONNEMENT

Déposez le fichier, choisissez un niveau. Utilisez la transcription.

Dans Voice Studio, cela prend quatre étapes. Via l'API REST et MCP, vous envoyez les fichiers audio directement depuis vos propres processus.

1

Ouvrir Voice Studio

Sous **Media AI** > **Studio vocal**, choisissez l'onglet **Transcription audio (STT)**. Depuis la page d'accueil, **Outils rapides** > **Média** > **Parole vers texte** y mène également.

2

Déposer le fichier

Faites glisser le fichier audio ou vidéo dans le champ. Laissez la langue sur Auto ou choisissez-la.

3

Choisir le niveau et les options

Sous **Qualité de transcription**, choisissez Basic, Premium ou Ultra. En Ultra, vous réglez le nombre de locuteurs sous **Reconnaissance des interlocuteurs**.

4

Utiliser la transcription

Copiez-la, utilisez **Télécharger au format TXT** pour l'enregistrer en fichier texte, ou transmettez-la à Translate Studio en un clic.

CARACTÉRISTIQUES TECHNIQUES

FONCTIONNALITÉ	DÉTAILS
Formats d'entrée	MP3, WAV, OGG, FLAC, M4A, MP4, WebM, AVI, MOV ; dans l'application jusqu'à 50 Mo, un fichier par lancement
Langues	Détection automatique ; 100 codes de langue via API, 15 choix fixes dans l'application
Niveaux de qualité	Basic (défaut API), Premium, Ultra (défaut de l'application)
Sortie via API	output_format json, srt et vtt à partir de Premium, rttm en Ultra
Séparation des locuteurs	En Ultra, min_speakers et max_speakers facultatifs, de 1 à 20
API REST	POST /job/add/paperoffice_voice___stt avec audio_file
MCP	po_stt_transcribe sur /mcp-full

03 · EN PRATIQUE

Réunion, dictée, vidéo de formation. Chacun en texte finalisé.

Compte rendu d'une réunion de projet

Le chef de projet téléverse l'enregistrement en Ultra, remplace Speaker 1 et Speaker 2 par leurs noms et copie le dialogue dans le compte rendu.

Dictée sur le terrain

Un géomètre dicte sur son téléphone sur la route et téléverse le fichier M4A au bureau. Basic renvoie le texte brut pour le rapport, et le vocabulaire connaît les noms de produits.

Sous-titres pour une vidéo de formation

Le développement envoie la vidéo via API avec `quality=premium` et `output_format=srt` et récupère le fichier de sous-titres avec horodatages.

QUESTIONS FRÉQUENTES

Quelles langues Voice STT reconnaît-il ?

La langue est détectée automatiquement. Via API, vous pouvez la définir parmi 100 codes de langue, et l'application propose 15 langues en choix fixes.

Voice STT peut-il distinguer les locuteurs ?

Oui, en Ultra. La transcription apparaît sous forme de dialogue, et les locuteurs peuvent être renommés. Via API, vous obtenez aussi le nombre de locuteurs, les chevauchements et le RTTM.

Puis-je obtenir des sous-titres en SRT ou VTT ?

Oui, via API à partir du niveau Premium avec `output_format=srt` ou `vtt`. Dans l'application, vous enregistrez la transcription en fichier TXT.

Comment la facturation est-elle établie ?

À l'usage, selon le niveau de qualité. Les tarifs actuels figurent sur la page des prix.

COMMENCER

Faites rédiger votre prochain enregistrement.

Connectez-vous à votre compte PaperOffice, ouvrez **Media AI > Studio vocal** et choisissez **Transcription audio (STT)**. Déposez le fichier, choisissez un niveau, lancez la transcription.

paperoffice.ai/fr/ai-voice-stt-reconnaissance-vocale-transcription-stt



EN SAVOIR PLUS SUR LE MODULE

paperoffice.ai/fr/ai-voice-stt-reconnaissance-vocale-transcription-stt

TARIFS

paperoffice.ai/fr/tarifs

CONTACT

paperoffice.ai/fr/contact