



DATENBLATT · ERSTELLEN · VOICE TTS

Jede Ansage neu ins Mikrofon sprechen? Ab jetzt spricht Ihr Text selbst.

Voice TTS macht aus Text natürlich klingende Sprache: über 800 Stimmen in 37 Sprachen, Dialoge mit mehreren Sprechern und Ihre eigene Stimme aus einer Sprachprobe. Für Ansagen, Anleitungen und Kundenkommunikation, in der App und per API.

AUF EINEN BLICK

Über 800 Stimmen mit Hörprobe, drei Qualitätsstufen, Dialoge mit Sprechermarken, Stimmenklonierung ab Premium und ein Aussprache-Lexikon für Namen und Marken.

01 · WORUM ES GEHT

Neu aufnehmen bei jeder Änderung? Text ändern, Audio neu erzeugen.

Die Telefonansage bekommt neue Öffnungszeiten, die Anleitung einen neuen Schritt, und wieder muss jemand ins Mikrofon. Mit Voice TTS ändern Sie nur den Text und erzeugen das Audio neu, in derselben Stimme und ohne Aufnahmetermin.

EINGABE

Text bis 50.000 Zeichen, für Dialoge mit einer Sprechermarke je Abschnitt

STIMMEN

Über 800 Stimmen in 37 Sprachen, weiblich und männlich, oder Ihre geklonte Stimme

ERGEBNIS

Audiodatei als MP3, ab Premium auch als WAV oder OGG

ZUGANG

App unter

Media AI › Voice Studio
, REST-API und MCP



Großer Stimmenkatalog

Über 800 Stimmen in 37 Sprachen, Spanisch und Portugiesisch in je zwei Varianten. In Basic stehen 5 Stimmen je Sprache bereit, ab Premium alle.

Dialoge mit mehreren Sprechern

Im Modus **Mehrere Sprecher** beginnt jeder Abschnitt mit Sprache und Stimme, etwa [de|Anneliese]. Das Studio fügt alle Abschnitte zu einem Audio zusammen.

Eigene Stimme

Unter **Stimmenklonierung** laden Sie eine Sprachprobe ab 10 Sekunden hoch, und Ihr Text wird in dieser Stimme gesprochen. Verfügbar ab Stufe Premium.

Aussprache-Lexikon

Im **TTS-Lexikon** legen Sie fest, wie Namen, Marken und Abkürzungen gesprochen werden. Die Einträge speichert Ihr Konto, ab Stufe Premium.

Drei Qualitätsstufen

Basic für die schnellste Erzeugung, Premium mit allen Stimmen, Ultra für die beste Qualität. Die Stufe wählen Sie unter **Modell**.

Sprechtempo

Das Tempo lässt sich fein einstellen, etwa 0,9 für etwas langsamer und 1,1 für etwas schneller. Standard ist 1,0.

02 · SO FUNKTIONIERT ES

Text eingeben, Stimme wählen. Audio anhören und laden.

Im Voice Studio sind es vier Schritte. Per REST-API und MCP stehen dieselben Stimmen, Stufen und Formate bereit.

1

Voice Studio öffnen

Unter **Media AI** › **Voice Studio** wählen Sie den Reiter **Text zu Sprache (TTS)** und geben Ihren Text ein oder fügen ihn ein.

2

Sprache, Stimme und Stufe wählen

Sprache und Stimme auswählen, unter **Modell** Basic, Premium oder Ultra einstellen. Für Dialoge schalten Sie auf **Mehrere Sprecher**.

3

Audio erzeugen

Das Studio erzeugt die Datei als Auftrag und zeigt sie zum Anhören an, sobald sie fertig ist.

4

Herunterladen und wiederverwenden

Mit **Herunterladen** speichern Sie die Datei. Im **Verlauf** laden Sie frühere Einstellungen und erzeugen das Audio neu.

TECHNISCHE DATEN

MERKMAL	DETAILS
Stimmen und Sprachen	über 800 Stimmen, 37 Sprachen, Liste mit Hörproben über GET /voice/voices
Textlänge	bis 50.000 Zeichen je Auftrag
Ausgabeformate	MP3, WAV, OGG; in der App WAV und OGG ab Stufe Premium
Sprachprobe für den Klon	WAV, MP3, M4A, FLAC, OGG, WebM, AAC; mindestens 10 Sekunden, bis 10 MB
Qualitätsstufen	Basic, Premium (Standard), Ultra
REST-API	POST /job/add/paperoffice_voice___tts, Dialoge mit voice=Multi
MCP	po_tts_generate, po_tts_multivoice_generate, po_tts_voice_clone, po_tts_voices_list

03 · IM EINSATZ

Hotline, Schulung, Produktvideo. Jeweils mit der passenden Stimme.

Neue Ansage für die Hotline

Der Support schreibt die Ansage, wählt eine Stimme und lädt das MP3 herunter. Ändern sich die Öffnungszeiten, passt er nur den Text an und erzeugt die Ansage neu.

Schulungsdialog mit zwei Stimmen

Die Personalabteilung schreibt einen Dialog mit den Marken [de|Anneliese] und [de|Florian] und erhält ein Audio, in dem beide abwechselnd sprechen.

Produktvideo in mehreren Sprachen

Das Marketing übersetzt das Skript im Übersetzungsstudio und vertont jede Fassung mit einer Stimme der Sprache. Das Lexikon sorgt für die richtige Aussprache des Produktnamens.

HÄUFIGE FRAGEN

Wie viele Stimmen und Sprachen gibt es?

Über 800 Stimmen in 37 Sprachen, Spanisch und Portugiesisch in je zwei Varianten. Die Liste mit Name, Geschlecht und Hörprobe liefert GET /voice/voices.

Wie entstehen Dialoge mit mehreren Sprechern?

Jeder Abschnitt beginnt mit einer Marke für Sprache und Stimme, etwa [en|Sarah]. Per API senden Sie den Text mit voice=Multi, jeder Abschnitt braucht dann eine Marke.

Kann ich meine eigene Stimme verwenden?

Ja, ab Stufe Premium. Sie laden eine Sprachprobe ab 10 Sekunden hoch, empfohlen sind 15 bis 30. Das Studio erkennt den gesprochenen Text der Probe, Sie können ihn korrigieren.

Wie wird abgerechnet?

Nach Verbrauch, abhängig von der Qualitätsstufe. Die aktuellen Preise stehen auf der Preisseite.

JETZT STARTEN

Erzeugen Sie Ihre erste Ansage im Voice Studio.

Melden Sie sich in Ihrem PaperOffice-Konto an, öffnen Sie **Media AI › Voice Studio** und wählen Sie **Text zu Sprache (TTS)**. Text eingeben, Stimme wählen, Audio erzeugen.

paperoffice.ai/de/ki-voice-tts-text-to-speech-sprachsynthese-tts



MEHR ZUM MODUL

paperoffice.ai/de/ki-voice-tts-text-to-speech-sprachsynthese-tts

PREISE

paperoffice.ai/de/preise

KONTAKT

paperoffice.ai/de/kontakt