



KARTA TECHNICZNA · PRZECHWYTYWANIE · MOWA NA TEKST (STT)

Spotkania przepisywać słowo w słowo? Od teraz każde nagranie wraca jako tekst.

Mowa na tekst zamienia nagrania spotkań, rozmów i dyktand w tekst. Sama rozpoznaje język, a poziom Ultra rozdziela mówców. Dla zespołów biurowych, sprzedaży i projektów, w aplikacji i przez API.

NA PIERWSZY RZUT OKA

Pliki audio i wideo do 50 MB, automatyczne rozpoznawanie języka, trzy poziomy jakości, rozdzielanie mówców w Ultra oraz znaczniki czasu słów, SRT i VTT przez API.

01 · O CZYM TO JEST

Najpierw spotkanie, potem pisanie? Teraz tekst przychodzi sam.

Protokół ze spotkania projektowego musi wyjść dziś, a nagranie trwa godzinę. W Mowie na tekst wgrywa się plik i otrzymuje tekst z podziałem na uczestników w Ultra, gotowy do skopiowania lub przetłumaczenia.

WEJŚCIE

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

JĘZYKI

Wykrywane automatycznie; 100 kodów języków przez API, 15 do
wyboru w aplikacji

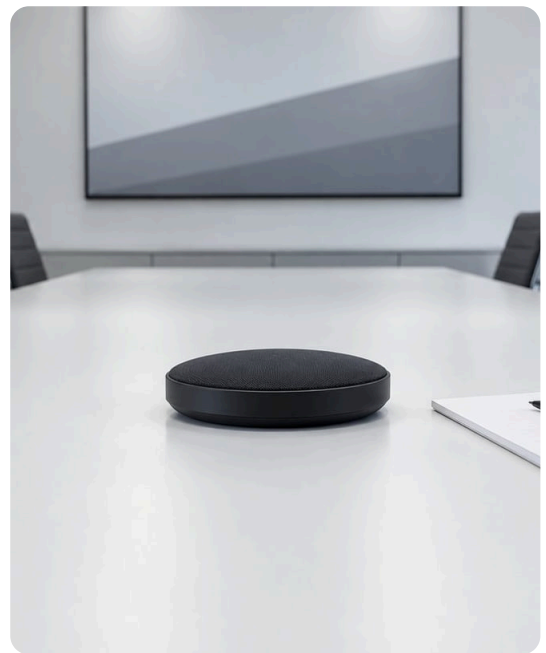
WYNIK

Transkrypcja jako zwykły tekst, w Ultra jako dialog z podziałem na
mówców; pobieranie jako TXT

DOSTĘP

Aplikacja w

Media AI › Voice Studio
, REST API i MCP

**Język wykrywany automatycznie**

Mowa na tekst sama rozpoznaje język nagrania. Jeśli go Państwo
znają, mogą go ustawić przez API, z 100 kodów języków.

Trzy poziomy jakości

Basic zwraca szybki zwykły tekst, Premium dodaje segmenty ze
znacznikami czasu słów, Ultra dodatkowo rozdziela mówców.
Poziom jest w polu **Jakość transkrypcji**.

Rozdział mówców

W Ultra Mowa na tekst rozpoznaje uczestników i pokazuje
transkrypcję jako dialog. Mówców można przemianować jednym
kliknięciem.

Słownictwo

Dzięki **Słownictwo specjalistyczne** Mowa na tekst dokładniej
rozpoznaje terminy fachowe z baz wiedzy i cenników. Przez API
dodają Państwo własne słowa priorytetowe od poziomu
Premium.

Napisy przez API

Przez API od poziomu Premium zwracane są SRT i VTT oraz
znaczniki czasu słów, a Ultra dodatkowo RTTM do analizy
mówców.

Dalej do Translate Studio

Użyj w **Translate Studio** przekazuje transkrypcję od razu do
Translate Studio, bez kopiowania i wklejania.

02 · JAK TO DZIAŁA

Wrzucić plik, wybrać poziom. Użyć transkrypcji.

W Voice Studio zajmuje to cztery kroki. Przez REST API i MCP wysyła się pliki audio bezpośrednio z własnych procesów.

1

Otworzyć Voice Studio

W **Media AI > Voice Studio** wybrać kartę **Zamiana mowy na tekst (STT)**. Ze strony startowej prowadzi tam też **Szybkie narzędzia > Media > Mowa na tekst**.

2

Wrzucić plik

Przeciagnąć plik audio lub wideo do pola. Język zostawić na Auto lub wybrać go ręcznie.

3

Wybrać poziom i opcje

W polu **Jakość transkrypcji** wybrać Basic, Premium lub Ultra. W Ultra liczbę mówców ustawia się w polu **Rozpoznawanie rozmówców**.

4

Użyć transkrypcji

Skopiować ją, zapisać jako plik tekstowy przyciskiem **Pobierz jako TXT** lub przekazać jednym kliknięciem do Translate Studio.

DANE TECHNICZNE

FUNKCJA	SZCZEGÓŁY
Formaty wejściowe	MP3, WAV, OGG, FLAC, M4A, MP4, WebM, AVI, MOV; w aplikacji do 50 MB, jeden plik na uruchomienie
Języki	automatyczne wykrywanie; 100 kodów języków przez API, 15 stałych wyborów w aplikacji
Poziomy jakości	Basic (domyślny w API), Premium, Ultra (domyślny w aplikacji)
Wyjście przez API	output_format json, srt i vtt od Premium, rttm w Ultra
Rozdział mówców	w Ultra opcjonalnie min_speakers i max_speakers od 1 do 20
REST API	POST /job/add/paperoffice_voice___stt z audio_file
MCP	po_stt_transcribe na /mcp-full

03 · W PRAKTYCE

Spotkanie, dyktando, film szkoleniowy. Każde jako gotowy tekst.

Protokół z projektowego spotkania

Kierownik projektu wgrzywa nagranie w Ultra, zmienia nazwy Mówca 1 i 2 na imiona uczestników i kopiuje dialog do protokołu.

Dyktat z terenu

Rzeczoznawca dyktuje telefonem w drodze, a w biurze wgrzywa plik M4A. Basic zwraca zwykły tekst do raportu, a słownictwo zna nazwy produktów.

Napisy do filmu szkoleniowego

Zespół programistów wysyła film przez API z `quality=premium` i `output_format=srt` i otrzymuje z powrotem plik napisów ze znacznikami czasu.

NAJCZĘŚCIEJ ZADAWANE PYTANIA

Jakie języki rozpoznaje Mowa na tekst?

Język wykrywany jest automatycznie. Przez API można go ustawić z 100 kodów języków, a aplikacja oferuje 15 języków jako stałe wybory.

Czy Mowa na tekst rozróżnia mówców?

Tak, w Ultra. Transkrypcja pojawia się jako dialog, a mówców można przemieszczać. Przez API dostaje się też liczbę mówców, nakładania się wypowiedzi i RTTM.

Czy mogę dostać napisy w SRT lub VTT?

Tak, przez API od poziomu Premium z `output_format=srt` lub `vtt`. W aplikacji transkrypcję zapisuje się jako plik TXT.

Jak to jest rozliczane?

Według zużycia, zależnie od poziomu jakości. Aktualne ceny podane są na stronie z cennikiem.

ROZPOCZNIJ

Niech kolejne nagranie zostanie spisane za Państwa.

Zalogować się na konto PaperOffice, otworzyć [Media AI > Voice Studio](#) i wybrać [Zamiana mowy na tekst \(STT\)](#). Wrzucić plik, wybrać poziom, uruchomić transkrypcję.

paperoffice.ai/pl/ai-voice-stt-mowa-na-tekst-transkrypcja-rozpoznawanie-stt



WIĘCEJ O MODULE

paperoffice.ai/pl/ai-voice-stt-mowa-na-tekst-transkrypcja-rozpoznawanie-stt

CENNIK

paperoffice.ai/en/pricing

KONTAKT

paperoffice.ai/pl/kontakt