

[AI & Machine Learning Products](https://cloud.google.com/products/machine-learning/) (<https://cloud.google.com/products/machine-learning/>)

[Cloud Speech-to-Text](https://cloud.google.com/speech-to-text/) (<https://cloud.google.com/speech-to-text/>)

[Documentation](https://cloud.google.com/speech-to-text/docs/) (<https://cloud.google.com/speech-to-text/docs/>) [Guides](#)

How-to guides

Performing speech recognition

Transcribing short audio files

Perform synchronous speech recognition on local and remote audio files.

(<https://cloud.google.com/speech-to-text/docs/sync-recognize>)

Transcribing long audio files

Perform asynchronous speech recognition on remote audio files.

(<https://cloud.google.com/speech-to-text/docs/async-recognize>)

Transcribing audio from streaming input

Perform streaming speech recognition on local audio files and other audio input streams.

(<https://cloud.google.com/speech-to-text/docs/streaming-recognize>)

Configuring recognition requests

Enabling word-level confidence

Specify that Cloud Speech-to-Text indicate a value of accuracy, or confidence level, for individual words in a transcription.

(<https://cloud.google.com/speech-to-text/docs/word-confidence>)

Separating different speakers

Get labels for different speakers in audio data transcribed by Cloud Speech-to-Text.

(<https://cloud.google.com/speech-to-text/docs/multiple-voices>)

Detecting language spoken automatically

Provide multiple language codes for audio transcription requests sent to Cloud Speech-to-Text.

(<https://cloud.google.com/speech-to-text/docs/multiple-languages>)

Transcribing audio with multiple channels

Transcribe audio files that include more than one channel.

(<https://cloud.google.com/speech-to-text/docs/multi-channel>)

Selecting a transcription model

Select a specialized machine learning model for audio transcription.

(<https://cloud.google.com/speech-to-text/docs/transcription-model>)

Enabling data logging

Enable data logging on your Google Cloud project to receive discounted pricing.

(<https://cloud.google.com/speech-to-text/docs/enable-data-logging>)

Using enhanced models

Use enhanced speech recognition models.

(<https://cloud.google.com/speech-to-text/docs/enhanced-models>)

Adding recognition metadata

Provide additional details about audio included in a transcription request to receive better output.

(<https://cloud.google.com/speech-to-text/docs/recognition-metadata>)

Getting punctuation

Include punctuation in transcription results from Cloud Speech-to-Text.

(<https://cloud.google.com/speech-to-text/docs/automatic-punctuation>)

Getting word timestamps

Perform speech recognition on a remote file and include time offset (timestamp) values for recognized words.

(<https://cloud.google.com/speech-to-text/docs/async-time-offsets>)

Base64 encoding

Learn how to base64 encode audio.

(<https://cloud.google.com/speech-to-text/docs/base64-encoding>)

Except as otherwise noted, the content of this page is licensed under the [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/) (<https://creativecommons.org/licenses/by/4.0/>), and code samples are licensed under the [Apache 2.0 License](https://www.apache.org/licenses/LICENSE-2.0) (<https://www.apache.org/licenses/LICENSE-2.0>). For details, see our [Site Policies](https://developers.google.com/terms/site-policies) (<https://developers.google.com/terms/site-policies>). Java is a registered trademark of Oracle and/or its affiliates.