General

What is Oracle Cloud Infrastructure Speech?

OCI Speech is a service that uses automatic speech recognition (ASR) to convert speech to text. The service enables developers, business units, content providers, tinkerers, and other users to transcribe audio files. With OCI Speech, users can transcribe call center calls or meetings, generate closed captions, and index and search audio and video content.

Why should I use OCI Speech?

You should use OCI Speech if you need a fast, accurate, time-stamped transcription service. If you are using OCI to store your audio files, you will also enjoy lower latencies and no network costs associated with transcription.

How do I get started with OCI Speech?

Start here to create your first transcription, or read more about the service here.

Features

What transcription services do you support?

We currently support file-based asynchronous transcription. We do not offer real-time transcription at this time.

What languages are currently supported?

Transcription comes with pretrained models for the following languages: English, Spanish, Portuguese, German, French, Italian and Hindi.

Are the files I transcribed used by OCI to improve the service (or for anything else)?

No. We only transcribe your content and keep no information from the file.

What else should I know about the service?

Like any other transcription service, the quality of the output depends on the quality of input audio file. Speakers' accents, background noises, switching between languages, using fusion languages (such as Spanglish), and multiple people speaking simultaneously can all impact the quality of transcription. We are constantly working to improve the performance of the service to provide more accurate transcriptions for all inputs and speakers.

Can OCI Speech automatically detect the language in the file?

Not currently (but soon).

What input file formats do you support?

We support single-channel, 16-bit PCM WAV audio files with a 16kHz sample rate. We also support the following media formats and will convert them to PCM WAV format before transcribing:

- AAC
- AC3
- AMR
- AU
- FLAC
- M4A
- MKV
- MP3
- MP4
- OGA
- OGG
- WAV
- WEBM

You can also perform conversion before submitting jobs to reduce latency. We recommend Audacity (GUI) or ffmpeg (command line) for audio transcoding.

What output formats do you support?

We support JSON (as default) and SRT (as the option with no further costs).

Billing and pricing

How will I be charged?

We use precision billing which means we charge you $0.50 for every hour of transcription, but we use seconds to measure the aggregated usage. For example, if you uploaded three files with the following durations: 10,860 seconds, 8,575 seconds, and 9,421 seconds, your monthly bill will by calculated by the sum of your seconds (28,856) divided by 3,600 (the number of seconds

in an hour) and minus 5 (number of free hours per month), multiplied by $0.50. In other words, you will be charged $1.508 or (28,856/3,600 - 5) x $0.50 = $1.508.

What is the billable metric for OCI Speech?

We named our billable metric "transcription hour." Transcription hour measures the number of audio hours transcribed during a given month of the service.

Are there any setup charges or minimum service commitments with Speech?

No. OCI Speech does not have any setup charges or minimum service commitments. And there's no hardware required.

Do you offer any free hours to try out the service?

Yes. We offer five hours of free transcription every month per tenancy.

Do you charge more for punctuation or SRT?

Punctuation is a free service just like SRT. Storing SRT files may increase your storage fee.

Other technical questions

What devices will be supported by OCI Speech?

Speech works with any recording device, and is not device-specific.

My file is not a WAV file. How should I convert my file to WAV?

We recommend using the ffmpeg utility with the following command: ffmpeg -i <input.ext> -fflags +bitexact -acodec pcm_s16le -ac 1 -ar 16000 <output.wav>.

I am getting the following error message: Either the bucket named "undefined" does not exist in the namespace <namespace> or you are not authorized to access it. How do I fix that?

See the [Speech Policy Setup](#).