Transcribing short audio files

This page demonstrates how to transcribe a short audio file to text using synchronous speech recognition.

Synchronous speech recognition returns the recognized text for short audio (less than ~1 minute) in the response as soon as it is processed. To process a speech recognition request for long audio, use Asynchronous Speech Recognition (https://cloud.google.com/speech-to-text/docs/async-recognize).

Audio content can be sent directly to Speech-to-Text, or it can process audio content that already resides in Google Cloud Storage. See also the audio limits (https://cloud.google.com/speech-to-text/quotas) for synchronous speech recognition requests.

The Cloud Speech-to-Text v1 is officially released and is generally available from the **`https://speech.googleapis.com/v1/speech`** endpoint. The Client Libraries (https://cloud.google.com/speech-to-text/docs/reference/libraries) are released as Alpha and will likely be changed in backward-incompatible ways. The client libraries are currently not recommended for production use.

These samples require that you have set up **`gcloud`** and have created and activated a service account. For information about setting up **`gcloud`**, and also creating and activating a service account, see Quickstart (https://cloud.google.com/speech-to-text/docs/quickstart).

Performing synchronous speech recognition on a local file

Here is an example of performing synchronous speech recognition on a local audio file:

PROTOCOL GCLOUD COMMAND MORE

Refer to the **`speech:recognize`** (https://cloud.google.com/speech-to-text/docs/reference/rest/v1/speech/recognize) API endpoint for complete details.

To perform synchronous speech recognition, make a `POST` request and provide the appropriate request body. The following shows an example of a `POST` request using `curl`. The example uses the access token for a service account set up for the project using the Google Cloud Cloud SDK (https://cloud.google.com/sdk). For instructions on installing the Cloud SDK, setting up a project with a service account, and obtaining an access token, see the quickstart (https://cloud.google.com/speech-to-text/docs/quickstart-protocol).

```
curl -X POST \
     -H "Authorization: Bearer "$(gcloud auth application-default print-access-toke
     -H "Content-Type: application/json; charset=utf-8" \
     --data "{
  'config': {
    'encoding': 'LINEAR16',
    'sampleRateHertz': 16000,
    'languageCode': 'en-US',
    'enableWordTimeOffsets': false
  },
  'audio': {
    'content': '/9j/7QBEUGhvdG9zaG9...base64-encoded-audio-content...fXNWzvDEeYxxxz
  }
}" "https://speech.googleapis.com/v1/speech:recognize"
```

See the RecognitionConfig (https://cloud.google.com/speech-to-text/docs/reference/rest/v1/RecognitionConfig) reference documentation for more information on configuring the request body.

The audio content supplied in the request body is base64-encoded. For more information on how to base64-encode audio, see Base64 Encoding Audio Content (https://cloud.google.com/speech-to-text/docs/base64-encoding). For more information on the `content` field, see RecognitionAudio (https://cloud.google.com/speech-to-text/docs/reference/rest/v1/RecognitionAudio).

If the request is successful, the server returns a `200 OK` HTTP status code and the response in JSON format:

```
{
  "results": [
    {
      "alternatives": [
        {
          "transcript": "how old is the Brooklyn Bridge",
          "confidence": 0.98267895
        }
      ]
    }
```

```
  ]
}
```

Performing synchronous speech recognition on a remote file

For your convenience, Speech-to-Text API can perform synchronous speech recognition directly on an audio file located in Google Cloud Storage, without the need to send the contents of the audio file in the body of your request.

Here is an example of performing synchronous speech recognition on a file located in Cloud Storage:

PROTOCOL GCLOUD COMMAND MORE

Refer to the `speech:recognize` (https://cloud.google.com/speech-to-text/docs/reference/rest/v1/speech/recognize) API endpoint for complete details.

To perform synchronous speech recognition, make a `POST` request and provide the appropriate request body. The following shows an example of a `POST` request using `curl`. The example uses the access token for a service account set up for the project using the Google Cloud Cloud SDK (https://cloud.google.com/sdk). For instructions on installing the Cloud SDK, setting up a project with a service account, and obtaining an access token, see the quickstart (https://cloud.google.com/speech-to-text/docs/quickstart-protocol).

```
curl -X POST -H "Authorization: Bearer "$(gcloud auth application-default print-acc
     -H "Content-Type: application/json; charset=utf-8" \
     --data "{
  'config': {
    'encoding': 'LINEAR16',
    'sampleRateHertz': 16000,
    'languageCode': 'en-US'
  },
  'audio': {
    'uri': 'gs://YOUR_BUCKET_NAME/YOUR_FILE_NAME'
  }
}" "https://speech.googleapis.com/v1/speech:recognize"
```

See the RecognitionConfig (https://cloud.google.com/speech-to-text/docs/reference/rest/v1/RecognitionConfig) reference documentation for more information on configuring the request body.

If the request is successful, the server returns a `200 OK` HTTP status code and the response in JSON format:

```
{
  "results": [
    {
      "alternatives": [
        {
          "transcript": "how old is the Brooklyn Bridge",
          "confidence": 0.98267895
        }
      ]
    }
  ]
}
```