Using enhanced models

This page describes how to request an enhanced speech recognition model when you send a transcription request to Speech-to-Text.

Google creates and improves *enhanced models* based upon data collected through data logging. If you opt-in to data logging (https://cloud.google.com/speech-to-text/docs/data-logging), you can help Google improve these models and also enjoy a discount on your usage.

There are currently two enhanced models: phone call and video. Using the improved phone call and video models, Cloud Speech-to-Text can more accurately recognize speech captured from these audio sources.

To use the enhanced recognition models, set the `useEnhanced` field to true, and then set the `model` field to your selected enhanced model in the `RecognitionConfig` (https://cloud.google.com/speech-to-text/docs/reference/rest/v1/RecognitionConfig) parameters for the request. Cloud Speech-to-Text supports enhanced models for all speech recognition methods: `speech:recognize` (https://cloud.google.com/speech-to-text/docs/reference/rest/v1/speech/recognize), `speech:longrunningrecognize` (https://cloud.google.com/speech-to-text/docs/reference/rest/v1/speech/longrunningrecognize), and `StreamingRecognizeRequest` (https://cloud.google.com/speech-to-text/docs/reference/rpc/google.cloud.speech.v1#google.cloud.speech.v1.StreamingRecognizeRequest).

The following code samples demonstrate how to request to use an enhanced model for a transcription request.

PROTOCOL C# GO 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 -s -H "Content-Type: application/json" \
    -H "Authorization: Bearer "$(gcloud auth application-default print-access-token
    https://speech.googleapis.com/v1/speech:recognize \
    --data "{
  'config': {
    'encoding': 'LINEAR16',
    'languageCode': 'en-US',
    'enableWordTimeOffsets': false,
    'enableAutomaticPunctuation': true,
    'model': 'phone_call',
    'useEnhanced': true
  },
  'audio': {
    'uri':'gs://cloud-samples-tests/speech/commercial_mono.wav'
  }
}"
```

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": "Hi, I'd like to buy a Chromecast. I was wondering whether
          "confidence": 0.8930228
        }
      ],
      "resultEndTime": "5.640s"
    },
    {
      "alternatives": [
        {
          "transcript": " Certainly, which color would you like? We are blue black
          "confidence": 0.9101991
        }
      ],
```

```
      "resultEndTime": "10.220s"
    },
    {
      "alternatives": [
        {
          "transcript": " Let's go with the black one.",
          "confidence": 0.8818244
        }
      ],
      "resultEndTime": "13.870s"
    },
    {
      "alternatives": [
        {
          "transcript": " Would you like the new Chromecast Ultra model or the regu
          "confidence": 0.94733626
        }
      ],
      "resultEndTime": "18.460s"
    },
    {
      "alternatives": [
        {
          "transcript": " Regular Chromecast is fine. Thank you. Okay. Sure. Would
          "confidence": 0.9519095
        }
      ],
      "resultEndTime": "25.930s"
    },
    {
      "alternatives": [
        {
          "transcript": " Express, please.",
          "confidence": 0.9101229
        }
      ],
      "resultEndTime": "28.260s"
    },
    {
      "alternatives": [
        {
          "transcript": " Terrific. It's on the way. Thank you. Thank you very much
          "confidence": 0.9321616
        }
      ],
```

```
      "resultEndTime": "34.150s"
    }
  ]
}
```

Caution: Enhanced models cost more than standard models. Review the pricing (https://cloud.google.com/speech-to-text/pricing) for more details.

What's next

- Review how to make synchronous transcription requests (https://cloud.google.com/speech-to-text/docs/sync-recognize).