This page shows how to detect text in an image, how to personalize translations, and how to generate synthetic speech from text. This tutorial uses Cloud Vision to detect text in an image file. Then, this tutorial shows how to use Cloud Translation to provide a custom translation of the detected text. Finally, this tutorial uses Text-to-Speech to provide machine dictation of the translated text.

1. Pass text recognized by the Cloud Vision API to the Cloud Translation API.
2. Create and use Cloud Translation glossaries to personalize Cloud Translation API translations.
3. Create an audio representation of translated text using the Text-to-Speech API.

Each Google Cloud API uses a separate pricing structure.

For pricing details, refer to the Cloud Vision pricing guide (/vision/pricing/), the Cloud Translation pricing guide (/translate/pricing), and the Text-to-Speech pricing guide (/text-to-speech/pricing).

Make sure that you have:

- A project in the Google Cloud Console (https://console.cloud.google.com/) with the Vision API, the Cloud Translation API, and the Text-to-Speech API enabled (/apis/docs/getting-started#enabling_apis)

- A basic familiarity with Python (https://www.python.org/) or NodeJS (https://nodejs.org/) programming

This tutorial uses Vision (/vision/docs/reference/libraries/), Translation (/translate/docs/reference/libraries/), and Text-to-Speech (/text-to-speech/docs/reference/libraries/) client libraries.

To install the relevant client libraries, run the following commands from the terminal.

Creating Translation glossaries requires using a service account key with "Cloud Translation API Editor" (/iam/docs/understanding-roles#cloud-translation-roles) permissions.

To set up a service account key with "Cloud Translation API Editor" permissions:

1. On the Google Cloud **Service accounts** page (https://console.cloud.google.com/iam-admin/serviceaccounts/), **Select a project**. Then, select **Create Service Account**. Designate the **Service account name** and **Create** the service account.
2. Under **Service account permissions**, click **Select a role**. Scroll to ***Cloud Translation*** and select ***Cloud Translation API Editor***. Select **Continue**.
3. Click **Create Key**, select **JSON**, and click **Create**.
4. In your terminal, set the ***GOOGLE_APPLICATION_CREDENTIALS*** variable using the following command. Replace ***path_to_key*** with the path to the downloaded JSON file containing your new service account key.

This tutorial uses the following system imports and client library imports.

You must associate a Google Cloud project (#prerequisites) with each request to a Google Cloud API. Designate your Google Cloud project (#prerequisites) by setting the `GCLOUD_PROJECT` environment variable from the terminal.

In the following command, replace ***project-id*** with your Google Cloud project ID. Run the following command from the terminal.

Use the Vision API to detect and extract text from an image. The Vision API uses Optical Character Recognition (OCR) (/vision/docs/ocr/) to support two text-detection features: detection of dense text, or `DOCUMENT_TEXT_DETECTION` (/vision/docs/features-list#text-detection), and sparse

text detection, or `TEXT_DETECTION`
(/vision/docs/features-list#document-text-detection-dense----------text--handwriting).

The following code shows how to use the Vision API `DOCUMENT_TEXT_DETECTION`
(/vision/docs/features-list#text-detection) feature to detect text in a photo with dense text.

After extracting text from an image, use Translation glossaries (/translate/docs/glossary) to personalize the translation of the extracted text. Glossaries provide pre-defined translations that override the Cloud Translation API translations of designated terms.

Glossary use cases include:

- **Product names:** For example, 'Google Home' must translate to 'Google Home'.
- **Ambiguous words:** For example, the word 'bat' can mean a piece of sports equipment or an animal. If you know that you are translating words about sports, you might want to use a glossary to feed the Cloud Translation API the sports translation of 'bat', not the animal translation.
- **Borrowed words:** For example, 'bouillabaisse' in French translates to 'bouillabaisse' in English; the English language borrowed the word 'bouillabaisse' from the French language. An English speaker lacking French cultural context might not know that bouillabaisse is a French fish stew dish. Glossaries can override a translation so that 'bouillabaisse' in French translates to 'fish stew' in English.

The Cloud Translation API accepts TSV, CSV, or TMX glossary files. This tutorial uses a CSV file uploaded to Cloud Storage to define sets of equivalent terms.

To make a glossary CSV file:

1. **Designate the language of a column** using either iso-639-1 (https://en.wikipedia.org/wiki/List_of_ISO_639-1_codes) or BCP-47 (https://tools.ietf.org/html/bcp47) language codes in the first row of the CSV file.

2. **List pairs of equivalent terms** in each row of the CSV file. Separate terms with commas. The following example defines the English translation for several culinary French words.

3. **Define variants of a word**. The Cloud Translation API is case-sensitive and sensitive to special characters such as accented words. Ensure that your glossary handles variations on a word by explicitly defining different spellings of the word.

4. Upload (/storage/docs/uploading-objects/) the glossary to a Cloud Storage bucket (/storage/docs/creating-buckets/). For the purposes of this tutorial, you do not need to upload a glossary file to a Cloud Storage bucket nor do you need to create a Cloud Storage bucket. Instead, use the publicly-available glossary file created for this tutorial to avoid incurring any Cloud Storage costs. Send the URI of a glossary file in Cloud Storage to the Cloud Translation API to create a glossary resource. The URI of the publicly-available glossary file for this tutorial is ***gs://cloud-samples-data/translation/bistro_glossary.csv*** (https://storage.cloud.google.com/cloud-samples-data/translation/bistro_glossary.csv). To download the glossary, click on the above URI link, but do not open it in a new tab.

In order to use a glossary, you must create a glossary resource with the Cloud Translation API. To create a glossary resource, send the URI of a glossary file in Cloud Storage to the Cloud Translation API.

Make sure that you are using a service account key with "Cloud Translation API Editor" (#setting_up_permissions_for_glossary_creation) permissions and make sure that you have set your project ID from the terminal (#setting_your_project_id).

The following function creates a glossary resource. With this glossary resource, you can personalize the translation request in the next step of this tutorial.

Once you create a glossary resource, you can use the glossary resource to personalize translations of text that you send to the Cloud Translation API.

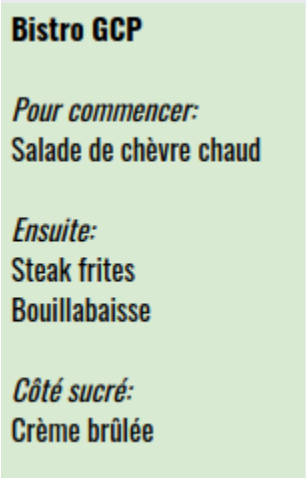
The following function uses your previously-created glossary resource to personalize the translation of text.

Now that you have personalized a translation of image-detected text, you are ready to use the Text-to-Speech API. The Text-to-Speech API can create synthetic audio of your translated text.

The Text-to-Speech API generates synthetic audio from either a string of plain text or a string of text marked up with Speech Synthesis Markup Language (SSML) (/text-to-speech/docs/ssml). SSML is a markup language which supports annotating text with SSML tags (/text-to-speech/docs/ssml#support-for-ssml-elements). You can use SSML tags to influence how the Text-to-Speech API formats synthetic speech creation (/text-to-speech/docs/ssml-tutorial).

The following function converts a string of SSML to an MP3 file of synthetic speech.

In the previous steps, you defined functions in `hybrid_glossaries.py` that use Vision, Translation, and Text-to-Speech. Now, you are ready to use these functions to generate synthetic speech of translated text from the following photo.



The following code calls functions defined in `hybrid_glossaries.py` to:

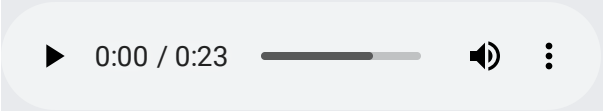
- create a Cloud Translation API glossary resource
- use the Vision API to detect text in the above image
- perform a Cloud Translation API glossary translation of the detected text
- generate Text-to-Speech synthetic speech of the translated text

To run the code, enter the following command in terminal in the directory (#downloading_the_code_samples) where your code is located:

The following output appears:

After running the code, navigate into the ***resources*** directory from the ***hybrid_glossaries*** directory. Check the resources directory for an `example.mp3` file.

Listen to the following audio clip to check that your `example.mp3` file sounds the same.



- Using a service account key without "Cloud Translation API Editor" permissions (#setting_up_permissions_for_glossary_creation) raises this exception.

- Not setting your GCLOUD_PROJECT variable (#setting_your_project_id) generates this error.

- Either using a glossary name which contains characters other than lowercase letters, digits, periods, a colon, or hyphens, or using a service account key without "Cloud Translation API Editor" permissions (#setting_up_permissions_for_glossary_creation) raises this exception.

- Setting the GOOGLE_APPLICATION_CREDENTIALS variable (#setting_up_permissions_for_glossary_creation) to an invalid filepath raises this exception.

- Not setting the GOOGLE_APPLICATION_CREDENTIALS variable (#setting_up_permissions_for_glossary_creation) raises this exception.

- Calling the Cloud Translation API, the Cloud Vision API, or the Text-to-Speech API without enabling their APIs (/apis/docs/getting-started#enabling_apis) generates this warning.

- Python 2.7.10 or earlier is not compatible with `HTML`. To fix this error, use a Python virtual environment (https://docs.python.org/3/library/venv.html). The virtual environment will use

the newest version of Python.

- Python 2.7.10 or earlier is not compatible with `HTML`. To fix this error, use a Python virtual environment (https://docs.python.org/3/library/venv.html). The virtual environment will use the newest version of Python.

Use the Google Cloud Console (https://console.cloud.google.com/) to delete your project if you do not need it. Deleting your project prevents incurring additional charges to your Google Cloud Platform account for the resources used in this tutorial.

1. In the Cloud Console (https://console.cloud.google.com/), go to the Projects page.
2. In the project list, select the project you want to delete and click **Delete**.
3. In the dialog box, type the project ID, and click **Shut down** to delete the project.

Congratulations! You just used Vision OCR to detect text in an image. Then, you created a Translation glossary and performed a translated with that glossary. Afterwards, you used Text-to-Speech to generate synthetic audio of the translated text.

To build on your knowledge of Vision, Translation, and Text-to-Speech:

- Make your own glossary. Learn how to create a Cloud Storage bucket (/storage/docs/creating-buckets/) and to upload your glossary CSV file to the bucket (/storage/docs/uploading-objects/).
- Experiment with other ways to use Translation glossaries (/translate/docs/glossary).

- Learn how to use Cloud Storage with Cloud Vision OCR (/functions/docs/tutorials/ocr/).
- Learn more about how to use SSML with Text-to-Speech (/text-to-speech/docs/ssml-tutorial/).
- Learn how to use the Vision API `imageContext` field (/vision/docs/handwriting#specify_the_language_optional) to pass along additional context about a photo when using Vision OCR.
- Explore community tutorials (/community/tutorials/).