

Image_Text-Speech Synthesizer

Vamshi Jaligama , Saikrishna Reddy Kotha , Pavan Kumar Ponugoti

I. Abstract

This document gives the detailed walk-through of the Image_Text-Speech Synthesizer project with the help of Google Cloud Platform

II. Introduction

We have implemented an image to text converter and then text to speech using the Cloud Vision API and Text to speech API in VM instances with the help of the google cloud platform .

III . Detailed Walkthrough

Firstly, we created a Project and then go to compute engine and create a instance with ubuntu OS.

Filter Enter property name or value

Status	Name	Zone	Recommendations	In use by	Internal IP	External IP	Connect
<input checked="" type="checkbox"/>	instance-1	us-central1-a			10.128.0.2 (nic0)	34.69.136.65 E2	SSH

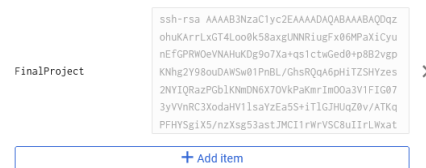
Then open the instance and click on edit and then we need to allow the https / https traffic in the instance firewall section

Firewalls
☒ Allow HTTP traffic
☒ Allow HTTPS traffic

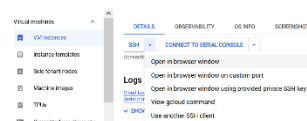
Network tags

http-server https-server

Next, we need to create a public key and then copy that public key into the VM instance by click on add SSH keys then copy and paste the public key into the box and click on save



After that open the SSH shell that is available in the VM instances in a new window



Enable the Cloud vision api for the Final project



vmshilgalga30@instance-1: ~ -- Mozilla Firefox

https://ssh.cloud.google.com/projects/finalproject-334307/zones/us-central-1-a/instances/nsta

Welcome to Ubuntu 20.04.3 LTS (GNU/Linux 5.11.0-1023-gcp x86_64)

- * Documentation: <https://help.ubuntu.com>
- * Management: <https://landscape.canonical.com>
- * Support: <https://ubuntu.com/advantage>

System information as of Sat Dec 11 03:27:59 UTC 2021

System load:	0.21	Processes:	117
Usage of /:	25.3% of 9.52GB	Users logged in:	1
Memory usage:	7%	IPv4 address for eno4:	10.128.0.2
Swap usage:	0%		

* Super-optimized for small spaces - read how we shrink the memory footprint of MicroK8s to make it the smallest full K8s around.

<https://ubuntu.com/blog/microk8s-memory-optimisation>

7 updates can be applied immediately.
To see these additional updates run: apt list --upgradable

Last login: Sat Dec 11 01:58:24 2021 from 35.235.245.129
vmshilgalga30@instance-1:~\$

Service accounts for project "RealProject"

A service account represents a Google Cloud service identity, such as code running on Compute Engine VMs, App Engine apps, or anytime running outside Google. [Learn more about service accounts.](#)

Organization policies can be used to secure service accounts and block risky account features, such as automatic IAM Grants, key creation/upload, or the creation of service accounts entirely. [Learn more about service account organization policies.](#)

Filter
Enter property name or value

Email	Status	Name	Description	Key ID	Key creation date	OAuth 2.0 Client ID	Actions
129748625727-computo@project-realproject.com	Active	Compute Engine default service account		00238236a323594dc2302c7a5ac0d81336a0	Dec 8, 2021	101421320170772801	<div> <div></div> <div></div> <div></div> </div>
final-5347619a@project-334307.com	Active	final		65a09915a8150a7194588a017a1a75fa10a7c5	Dec 8, 2021	1096964704230268	<div> <div></div> <div></div> <div></div> </div>

DETAILS
PERMISSIONS
KEYS
METRICS
LOGS

Keys

Service account keys could pose a security risk if compromised. We recommend you avoid downloading service account keys and instead use the [Download Identity Federation](#). You can learn more about the best way to authenticate service accounts on Google Cloud [here](#).

Add a new key or upload a public key certificate from an existing key pair.

Block service account key creation using [organization policies](#).
[Learn more about setting organization policies for service accounts](#)

ADD KEY
+

Type	Status	Key	Key creation date	Key expiration date
Active	Active	65a09915a8150a7194588a017a1a75fa10a7c5	Dec 8, 2021	Dec 31, 9999

```
7 updates can be applied immediately.
to see these additional updates run: apt list --upgradable

Last login: Sat Dec 11 01:58:24 2021 from 35.235.245.129
[rm@hij1alligama309instance:~]$ cd /final
bash: cd: /final: No such file or directory
[rm@hij1alligama309instance:~]$ cd /final
[rm@hij1alligama309instance:~]$ ./Final5_VI finalprotect-334307-002388266e32_1.iso
```

[illegible]

```
import io
from google.cloud import vision
from google.cloud.vision_v1 import types

vision = vision.ImageAnnotatorClient()

image_value = 'albert-einstein-quotes-01-scaled.jpg'

with io.open(image_value, 'rb') as files:
    content = files.read()
image = types.Image(content=content)

value_response = vision.text_detection(image = image)

image_to_text = value_response.text_annotations

output_text = open('quotes.txt', 'w')

for texts in image_to_text:

    print(texts.description)

    print(texts.description, file = output_text)

output_text.close()
```

Here is python code for the conversion of the text into the speech:

```
from google.cloud import texttospeech
import os

text_to_speech = texttospeech.TextToSpeechClient()

image_text = 'quotes.txt'

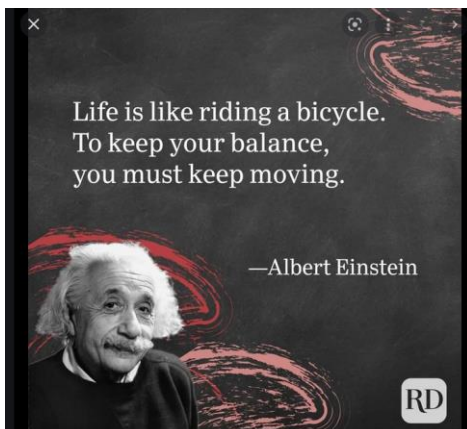
with open(image_text, 'r') as files:
    text = files.read()
    value = texttospeech.types.SynthesisInput(text=text)

speech = texttospeech.types.VoiceSelectionParams(language_code='en-US',
                                                  ssml_gender=texttospeech.enums.SsmlVoiceGender.MALE)
configuration = texttospeech.types.AudioConfig(audio_encoding=texttospeech.enums.AudioEncoding.MP3)
result = text_to_speech.synthesize_speech(value, speech, configuration)

with open('file_output_speech1.mp3', 'wb') as final_speech:
    final_speech.write(result.audio_content)
```

Then we need to search for a image of our choice after that we need to copy that image link and go to the shell there, by using the wget command and download it to the instance.

Google search image



Wget command for download:

```
root@kali:~/workspace# wget https://www.rd.com/wp-content/uploads/2018/04/Albert-Einstein-quotes-01-scaled.jpg
--2023-12-21 02:17:01-- https://www.rd.com/wp-content/uploads/2018/04/Albert-Einstein-quotes-01-scaled.jpg
Resolving www.rd.com (www.rd.com)... 104.18.23.214, 104.18.23.214, 104.18.23.214, 104.18.23.214, 104.18.23.214, ...
Connecting to www.rd.com (www.rd.com) [104.18.23.214]:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 48153 (47K) [image/jpeg]
Saving to: 'Albert-Einstein-quotes-01-scaled.jpg'

Albert-Einstein-quotes-01-scaled.jpg 100% [48153/48153] 100KB/s 0s 0.00s

2023-12-21 02:17:01 (10.3 MB/s) - 'Albert-Einstein-quotes-01-scaled.jpg' saved [48153/48153]
```

After downloading the image file then we need to look for the name of the image file and then copy that file name in the converter python code.

```
vision = vision.ImageAnnotatorClient()

image_value = 'albert-einstein-quotes-01-scaled.jpg'
```

Next we need to run the python code for converting the image to text :

```
root@kali:~/workspace# python3 converter2.py
Life is like riding a bicycle.
To keep your balance,
you must keep moving.
—Albert Einstein
RD

Life
is
like
riding
a
bicycle.
To
keep
your
balance,
you
must
keep
moving.
—Albert
Einstein
```

After running the text to speech python code a new mp3 file is created in the instance we can see that a new file created called file_output_speech1.mp3

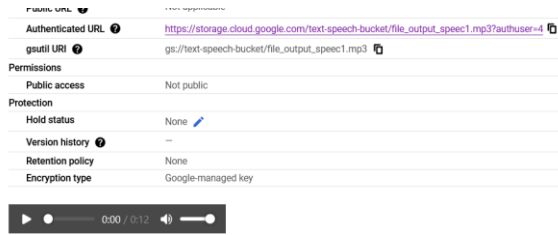
```
root@kali:~/workspace# ls
albert-einstein-quotes-01-scaled.jpg  converter2.py  converter3.py  file_output_speech1.mp3  finalproject-334807-fdd01833fe95.json  quotes.txt  testspeech1.py
root@kali:~/workspace#
```

Finally we need to upload this mp3 into the bucket so that we can access the mp3 file in gcp using the command “gsutil cp file_output_speech1.mp3 gs://text-speech-bucket”

```
root@kali:~/workspace# gsutil cp finalproject-334807-fdd01833fe95.json gs://text-speech-bucket
Copying file://finalproject-334807-fdd01833fe95.json [Content-Type=application/json]...
/ [1 files] 2.3 KiB/ 2.3 KiB
Operation completed over 1 objects/2.3 KiB.
```

We can see that a mp3 is uploaded in the bucket

OBJECTS	CONFIGURATION	PERMISSIONS	PROTECTION	LIFECYCLE						
Buckets > test-speech-bucket										
UPLOAD FILES	UPLOAD FOLDER	CREATE FOLDER	MANAGE HOLDS	DOWNLOAD						
Filter by name prefix only										
Filter objects and folders										
Show deleted data										
<input type="checkbox"/>	Name	Size	Type	Created	Storage class	Last modified	Public access	Version history	Encryption	Retention
<input type="checkbox"/>	file_output_speech1.mp3	48.15 KiB	audio/mpeg	Dec 10, 2023	Standard	Dec 10, 2023	Not public	—	Google-managed only	2



References

- [1] GCP, "Cloud Vision Quick start," [Online]. Available:
<https://cloud.google.com/vision/docs/samples>.
- [2] GCP, "Cloud Text to Speech Quickstart," [Online]. Available:
https://cloud.google.com/text-to-speech/docs/quickstart-protocol?hl=en_US.