



2020

# Voice to text

There was a young flower in the desert where all was dry and sad looking.  
It was growing by itself . enjoy every day



- 01. 語音介紹
- 02. Speech Recognition
- 03. 程式碼



PART ONE

## 語言介紹

關於語言辨識



## 介紹

### 語音辨識

語音辨識（speech recognition）技術，也被稱為自動語音辨識（英語：Automatic Speech Recognition, ASR）、電腦語音識別或是語音轉文字識別（英語：Speech To Text, STT），其目標是以電腦自動將人類的語音內容轉換為相應的文字。與說話人辨識及說話人確認不同，後者嘗試辨識或確認發出語音的說話人而非其中所包含的詞彙內容。





## 介紹

Please click here to add the required titles

## 語音辨識

語音辨識技術的應用包括語音撥號、語音導航、室內裝置控制、語音文件檢索、簡單的聽寫資料錄入等。語音辨識技術與其他自然語言處理技術如機器翻譯及語音合成技術相結合，可以構建出更加複雜的應用，例如語音到語音的翻譯。

語音辨識技術所涉及的領域包括：

- 訊號處理
- 圖型識別
- 概率論和資訊理論
- 發聲機理和聽覺機理
- 人工智慧、文本探勘等等。



PART TWO

Speech Recognition



## Speech Recognition



### Voice to Text

處理語音轉文字的常見API，如Google Speech、IBM Speech to Text、Microsoft Bing Speech、AT&T Speech。其中Google 及 Microsoft在語音轉文字的開發語言為Python，對於後面所需的Text Mining會很方便。而Microsoft Bing 對於採樣頻率要求較麻煩，因此我們採用Google Speech作為測試。



# Speech Recognition

## python函式庫安裝

使用Google Cloud Speech-to-text就需要API Key

雖然準確率不低、轉換速度也快，但是要價不斐且申請麻煩

<https://cloud.google.com/speech-to-text/>

功能	標準級模型 (強化版通話和影片模型以外的所有模型)		進階級模型* (強化版通話和影片模型)	
	0 至 60 分鐘	61 分鐘以上，100 萬分鐘以下	0 至 60 分鐘	61 分鐘以上，100 萬分鐘以下
語音辨識 (未啟用資料記錄功能 - 預設值)	免費	每 15 秒 \$0.006 美元**	免費	每 15 秒 \$0.009 美元**
語音辨識 (已啟用資料記錄功能)	免費	每 15 秒 \$0.004 美元**	免費	每 15 秒 \$0.006 美元**

因此我們使用Google 的套件

輸入pip install speechRecognition

或是 conda install speechRecognition

```
命令提示字元
Microsoft Windows [版本 10.0.18362.476]
(c) 2019 Microsoft Corporation. 著作權所有，並保留一切權利。
C:\Users\ADS>pip install speechRecognition
Requirement already satisfied: speechRecognition in d:\anaconda\lib\site-packages (3.8.1)
```



PART THREE

## 發展現狀

Click here to add a little title

點擊此處添加文字描述，文字顏色、大小、字體可根據需求  
進行更改。



## Voice to text

由 音 檔 轉 文 字

```
import speech_recognition  
r = speech_recognition.Recognizer()  
with speech_recognition.AudioFile("data1_new.wav") as source:  
    audio = r.record(source)  
    text = r.recognize_google(audio,language='en-US') → 中文 : zh-TW  
    英文 : en-US  
print(text)
```

The screenshot shows a Windows Command Prompt window titled '命令提示字元' (Command Prompt). The command entered is 'python audioToText.py'. The output displayed is a single line of Chinese text: '我有一隻小毛驥我從來也不騎有一天我心血來潮騎著去趕集我手裡拿著小皮鞭我的心裡好個理由'. This demonstrates that the script successfully converted the audio file 'data1\_new.wav' into text, using the 'en-US' language parameter which was intended to be overridden by the Chinese output.

使用的檔案



## Voice to text

由麥克風轉文字

```
import speech_recognition as sr
```

使用的檔案

```
r=sr.Recognizer()  
with sr.Microphone() as source:  
    print("Please wait. Calibrating microphone...")
```

```
r.adjust_for_ambient_noise(source, duration=5) —————→ 錄音時間  
print("Say something!")  
audio=r.listen(source)
```

try:

```
    print("Google Speech Recognition thinks you said:")  
    print(r.recognize_google(audio, language="zh-TW"))  
except sr.UnknownValueError:  
    print("Google Speech Recognition could not understand audio")  
except sr.RequestError as e:  
    print("No response from Google Speech Recognition service: {0}".format(e))
```



# Voice to text

輸出文檔

```
f = open('a.txt','w')  
f.write(text)  
f.close()
```

開啟的檔案及權限，此權限會清空原本的內文

寫入文字並釋放記憶體  
若未釋放，則文字不會寫入完成





# THE END