

Internet de las Cosas

# Prácticas

Reconocimiento de Voz



Rogelio Ferreira Escutia

Profesor / Investigador  
Tecnológico Nacional de México  
Campus Morelia



# Reconocimiento de Voz

# Reconocimiento de Voz

- **Procesos**



# Reconocimiento de Voz

- **Clasificación de Motores de Voz:**
- **Online (en línea).**
  - **At&T**
  - **Wit.ai**
  - **IBM**
  - **Google**
- **Offline (fuera de línea)**
  - **Julius.**
  - **Sphinx**

# Google

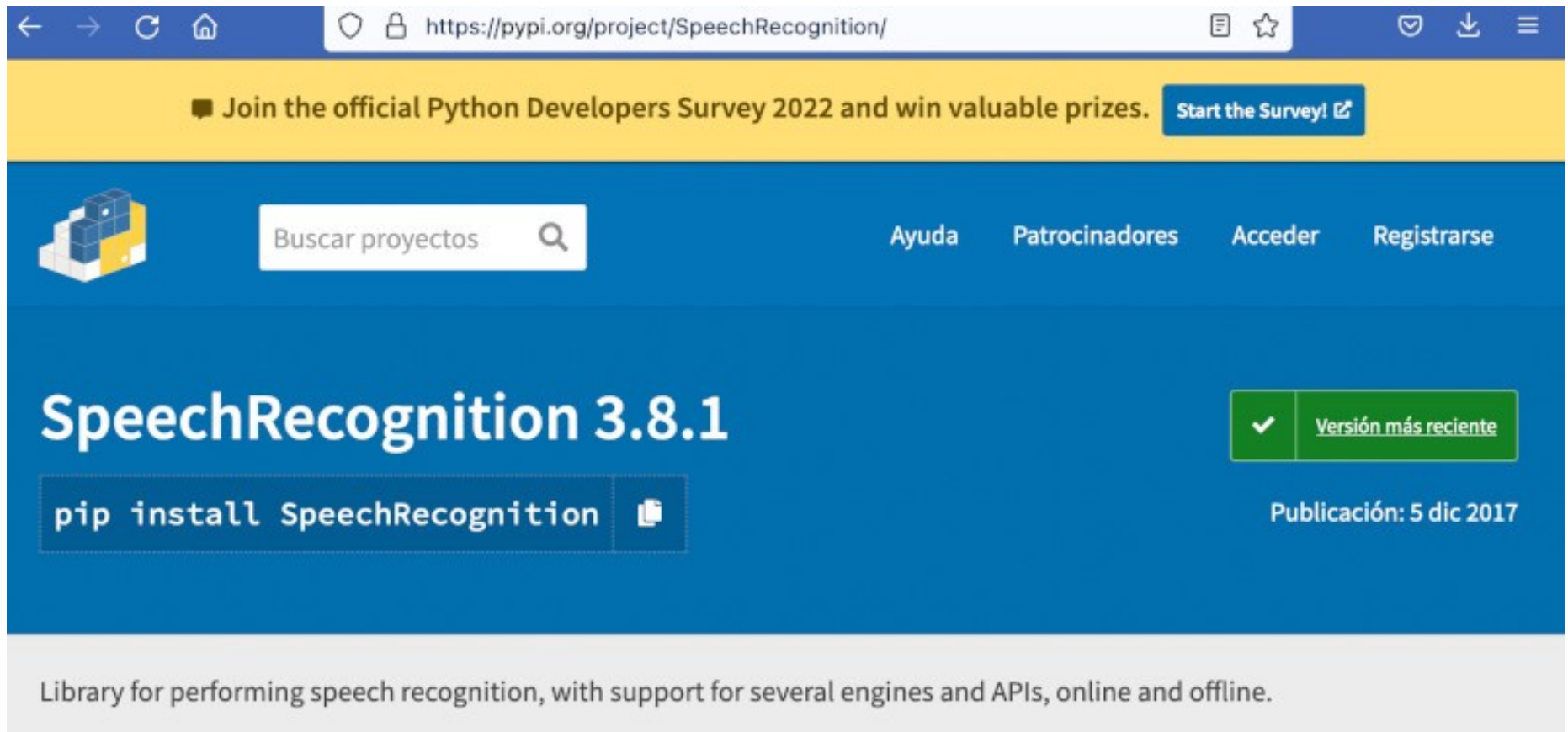
## ■ Precios:

Función	Modelos estándar (todos excepto los modelos mejorados para vídeo y llamadas telefónicas)		Modelos mejorados (para vídeo y llamadas telefónicas)	
	De 0 a 60 minutos	De más de 60 minutos a 1 millón de minutos	De 0 a 60 minutos	De más de 60 minutos a 1 millón de minutos
Reconocimiento de voz (opción predeterminada, sin almacenamiento de registros de datos)	Gratis	0,006 USD por cada 15 segundos**	Gratis	0,009 USD por cada 15 segundos**
Reconocimiento de voz (con el almacenamiento de registros de datos habilitado)	Gratis	0,004 USD por cada 15 segundos**	Gratis	0,006 USD por cada 15 segundos**

Voz a Texto

# Python

- **SpeechRecognition:**



The screenshot shows the PyPI project page for SpeechRecognition. At the top, there is a navigation bar with a search box labeled "Buscar proyectos" and links for "Ayuda", "Patrocinadores", "Acceder", and "Registrarse". A yellow banner at the top of the page promotes the "Python Developers Survey 2022". The main content area features the project name "SpeechRecognition 3.8.1" in large white text on a blue background. Below the name is a code block containing the command `pip install SpeechRecognition` and a copy icon. To the right of the name, there is a green badge with a checkmark and the text "Versión más reciente", and below that, the publication date "Publicación: 5 dic 2017". At the bottom of the page, a grey box contains the description: "Library for performing speech recognition, with support for several engines and APIs, online and offline."

```
# python_voz_microfono_texto.py
#
# https://pypi.org/project/SpeechRecognition/
# pip3 install SpeechRecognition

# https://pypi.org/project/PyAudio/
# apt install python3-pyaudio
#
# apt install flac

import speech_recognition as sr
import pyaudio

voz = sr.Recognizer()

print("\nEscuchando...")

with sr.Microphone() as fuente:
    voz.adjust_for_ambient_noise(fuente)
    audio = voz.listen(fuente)
    texto = voz.recognize_google(audio)

print("\nTexto: ", texto)
```





```
1 # https://pypi.org/project/SpeechRecognition/
2 # pip3 install SpeechRecognition
3
4 # https://pypi.org/project/PyAudio/
5 # apt install python3-pyaudio
6 #
7 # apt install flac
8
9 import speech_recognition as sr
10 import pyaudio
11
12 voz = sr.Recognizer()
13
14 print("\nEscuchando...")
15
16 with sr.Microphone() as fuente:
17     voz.adjust_for_ambient_noise(fuente)
18     audio = voz.listen(fuente)
19     texto = voz.recognize_google(audio)
20
21 print("\nTexto: ", texto)
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL JUPYTER

Python + v □

```
ALSA lib conf.c:4745:(_snd_config_evaluate) function snd_func_refer returned error: No such file or directory
ALSA lib conf.c:5233:(snd_config_expand) Evaluate error: No such file or directory
ALSA lib pcm.c:2660:(snd_pcm_open_noupdate) Unknown PCM iec958:{AES0 0x6 AES1 0x82 AES2 0x0 AES3 0x2 CARD 0}
ALSA lib pcm_usb_stream.c:486:(_snd_pcm_usb_stream_open) Invalid type for card
ALSA lib pcm_usb_stream.c:486:(_snd_pcm_usb_stream_open) Invalid type for card
Cannot connect to server socket err = No such file or directory
Cannot connect to server request channel
jack server is not running or cannot be started
JackShmReadWritePtr::~JackShmReadWritePtr - Init not done for -1, skipping unlock
JackShmReadWritePtr::~JackShmReadWritePtr - Init not done for -1, skipping unlock
```

Texto: hola Buenas Tardes como estas

rogelio@raspberrypi:~/python \$

```
# https://pypi.org/project/SpeechRecognition/
# pip3 install SpeechRecognition
# https://pypi.org/project/PyAudio/
# apt install python3-pyaudio
# apt install flac
# Página de lenguajes soportados:
# https://stackoverflow.com/questions/14257598/what-are-language-codes-in-chromes-implementation-of-the-html5-speech-recogniti/14302134#14302134
import speech_recognition as sr
import pyaudio
voz = sr.Recognizer()
while True:
    print("\nPython escuchando...")
    texto = ""
    with sr.Microphone() as fuente:
        voz.adjust_for_ambient_noise(fuente)
        try:
            audio = voz.listen(fuente)
            texto = voz.recognize_google(audio, language="es-MX")
        except sr.UnknownValueError:
            pass
        except sr.RequestError:
            pass
        except sr.WaitTimeoutError:
            pass
    print("\nDijiste: ", texto)
    if texto == "adiós":
        print("\nHasta la vista Beiby!\n")
        break
```



```

12 import speech_recognition as sr
13 import pyaudio
14
15 voz = sr.Recognizer()
16
17 while True:
18     print("\nPython escuchando...")
19     texto = ""
20     with sr.Microphone() as fuente:
21         voz.adjust_for_ambient_noise(fuente)
22         try:
23             audio = voz.listen(fuente)
24             texto = voz.recognize_google(audio, language="es-MX")
25         except sr.UnknownValueError:
26             pass
27         except sr.RequestError:
28             pass
29         except sr.WaitTimeoutError:
30             pass
31     print("\nDijiste: ", texto)
32     if texto == "adios":
33         print("Hasta la vista Beiby!")
34         break

```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL JUPYTER

```

Dijiste:
Python escuchando...
Dijiste: ahora sí
Python escuchando...
Dijiste: A qué hora vas al pan
Python escuchando...

```



# Wit.ai

```
# recognize speech using Wit.ai
WIT_AI_KEY = "INSERT WIT.AI API KEY HERE"

try:
    print("Wit.ai thinks you said " + r.recognize_wit(audio, key=WIT_AI_KEY))
except sr.UnknownValueError:
    print("Wit.ai could not understand audio")
except sr.RequestError as e:
    print("Could not request results from Wit.ai service; {0}".format(e))
```



# IBM

```
# recognize speech using IBM Speech to Text
IBM_USERNAME = "INSERT IBM SPEECH TO TEXT USERNAME HERE"
IBM_PASSWORD = "INSERT IBM SPEECH TO TEXT PASSWORD HERE"

try:
    print("IBM Speech to Text thinks you said " + r.recognize_ibm(audio,
        username=IBM_USERNAME, password=IBM_PASSWORD))
except sr.UnknownValueError:
    print("IBM Speech to Text could not understand audio")
except sr.RequestError as e:
    print("Could not request results from IBM Speech to Text service;
        {0}".format(e))
```





# Reconocimiento de Voz

- Usando directamente el micrófono:

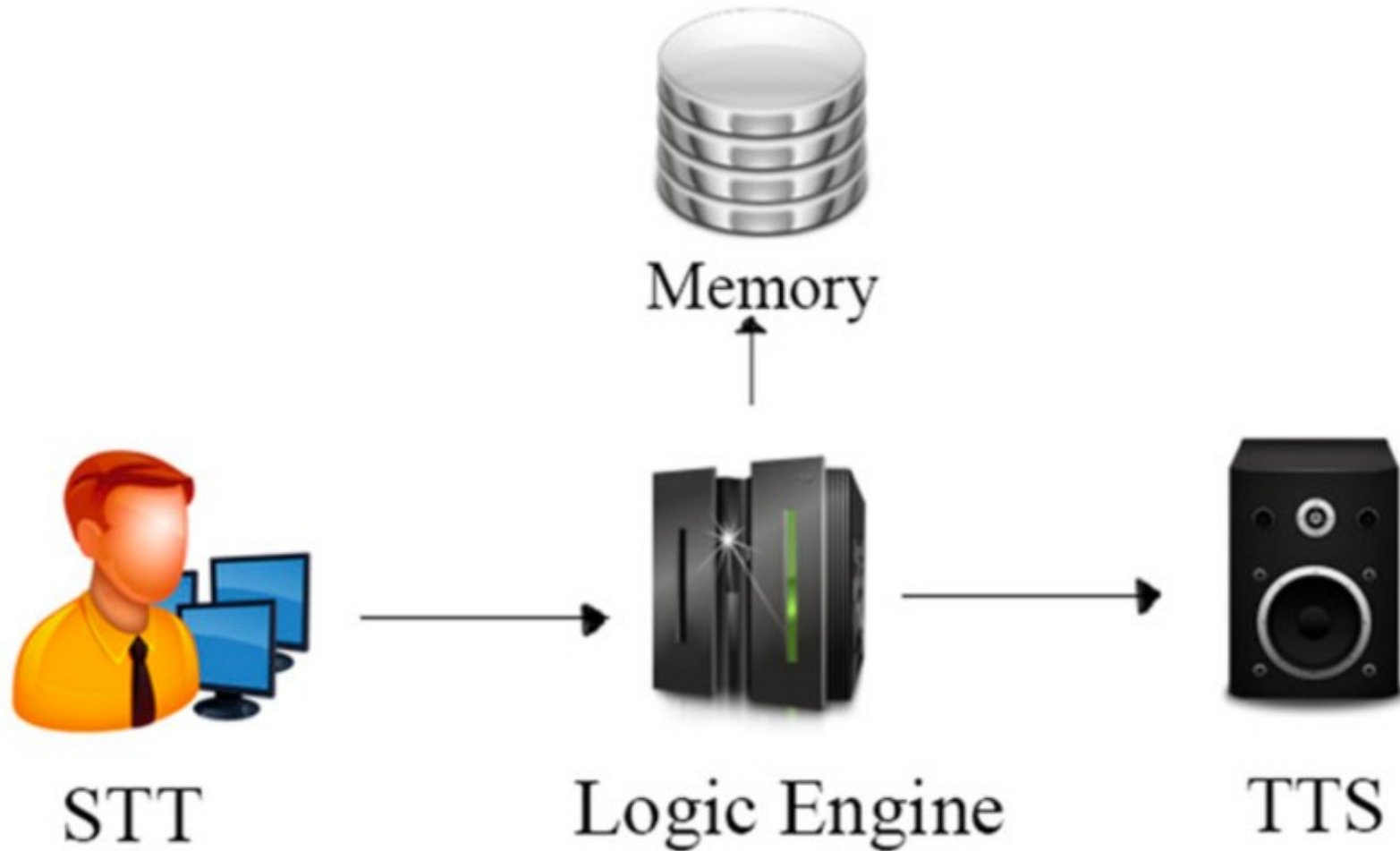
```
import speech_recognition as sr
speech = sr.Recognizer()
print('Python is listening...')
with sr.Microphone() as source:
    speech.adjust_for_ambient_noise(source)
    audio = speech.listen(source)
    inp = speech.recognize_google(audio)
print(f'You just said {inp}.')
```

# Análisis del Mensaje

(Procesamiento del  
Lenguaje Natural)

# Análisis del Mensaje

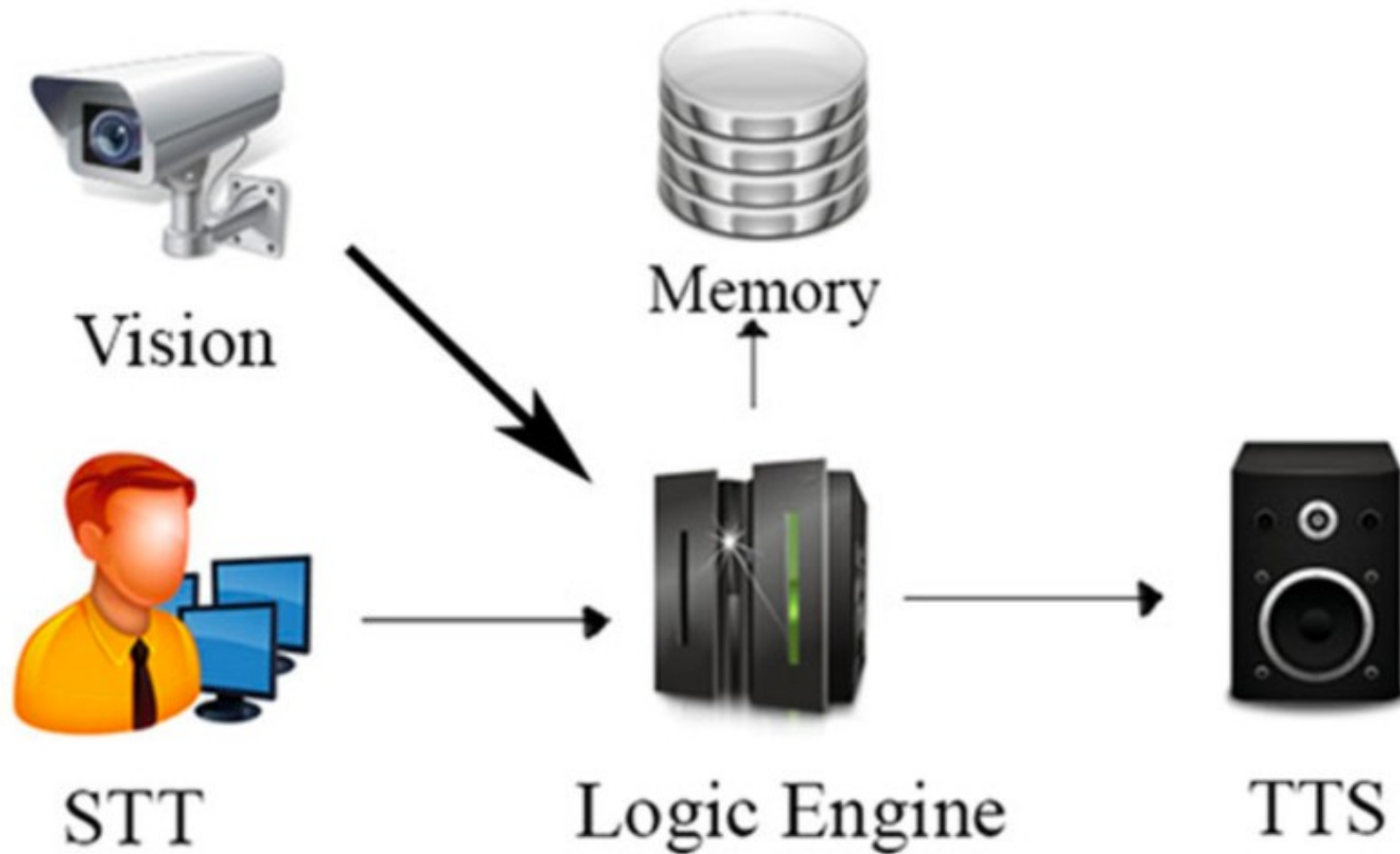
- **Procesamiento del Lenguaje Natural:**





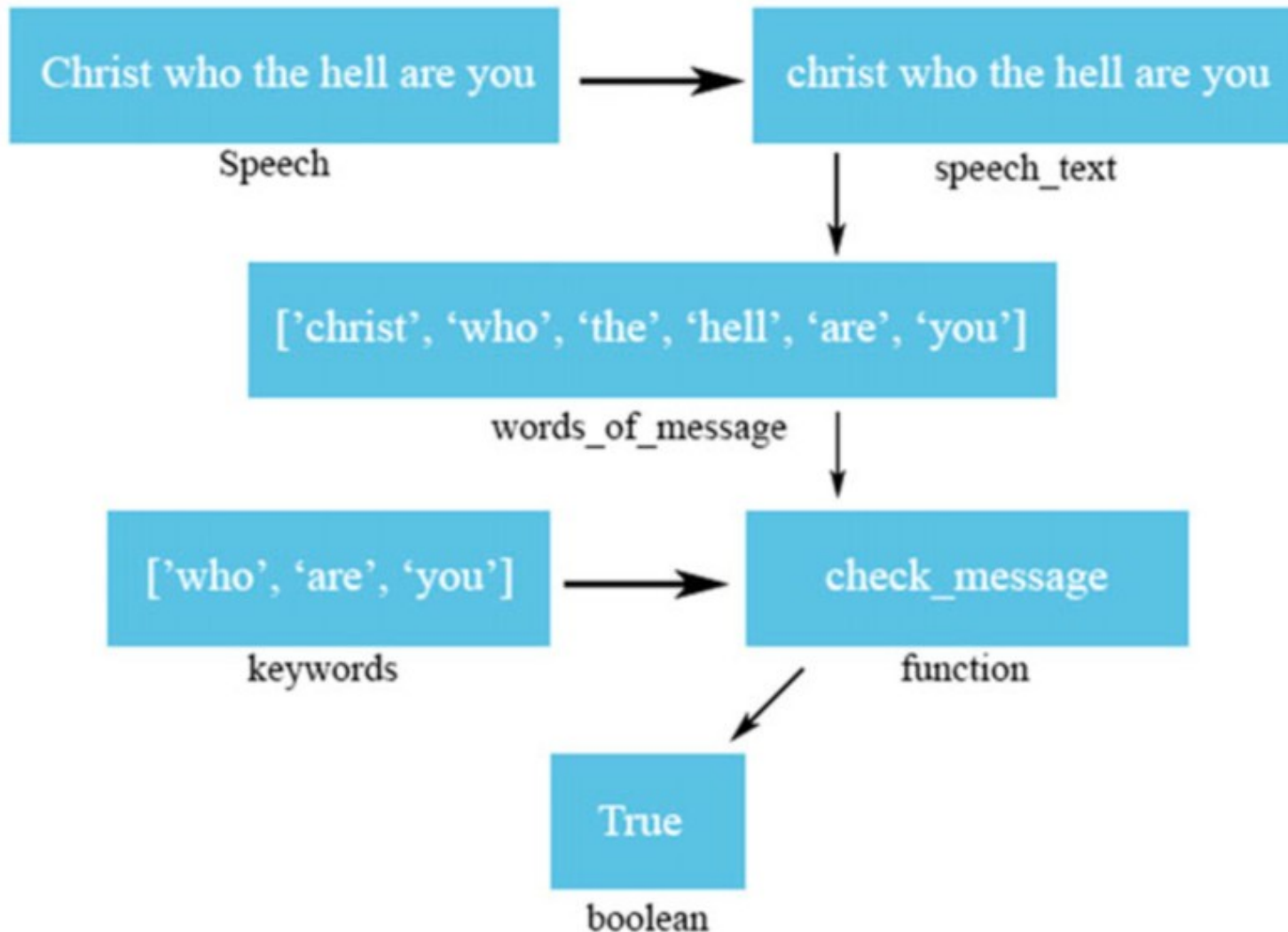
# Análisis del Mensaje

- **Procesamiento del Lenguaje Natural:**



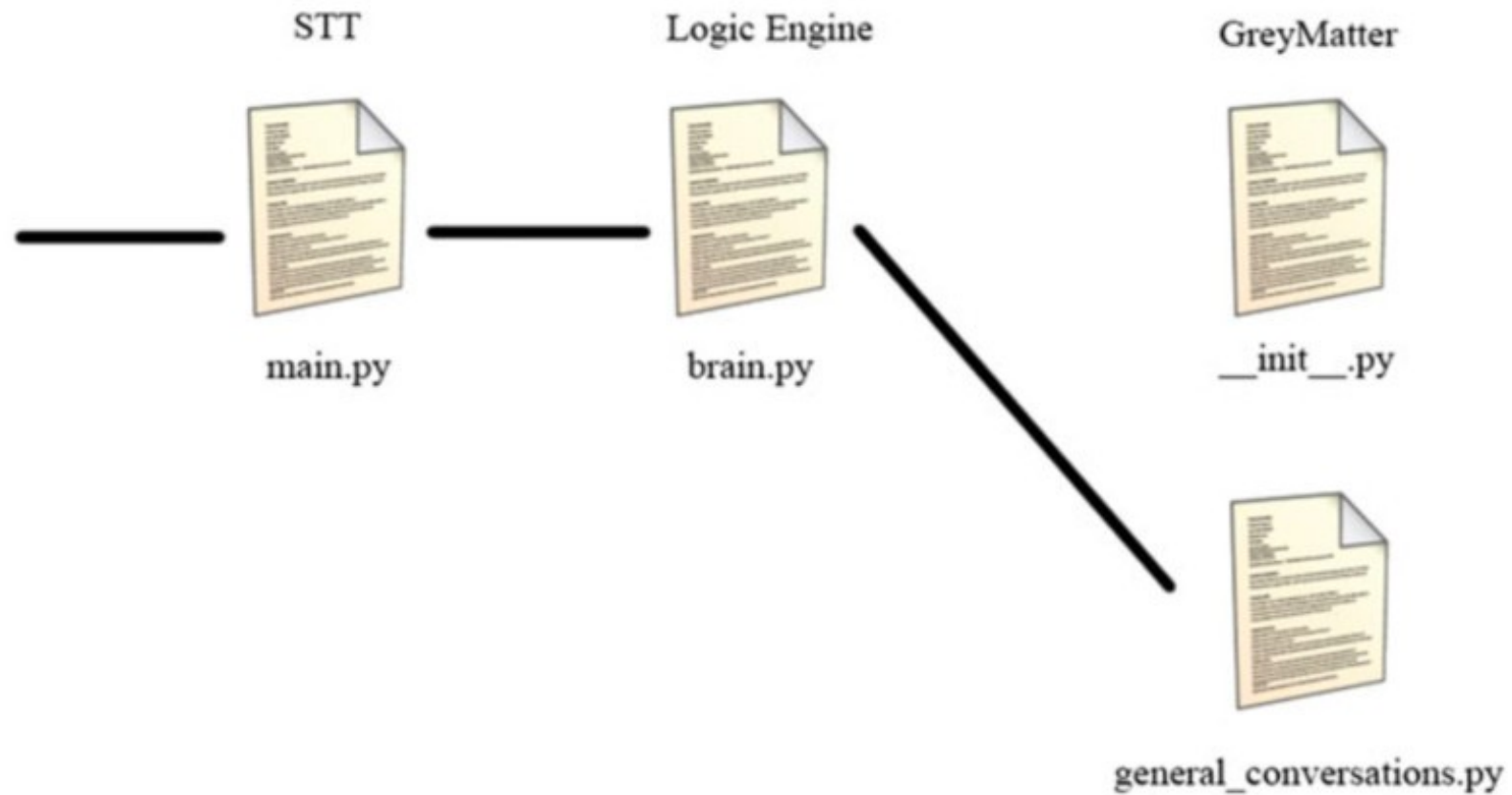
# Análisis del Mensaje

- **Procesamiento del Lenguaje Natural:**



# Análisis del Mensaje

- **Procesamiento del Lenguaje Natural:**



Texto a Voz

# Generación de Voz (sonido)

```
import os
from playsound import playsound
os.system('gtts-cli --nocheck "Hola, cómo estás?" | mpg123 -q -')
from gtts import gTTS
tts = gTTS('Hola mundo. Estamos convirtiendo texto a voz con Python!', lang='es-us')
tts.save("voz.mp3")
os.system("mpg123 " + "voz.mp3")
```

# Generación de Voz (sonido)

```
import os
import sys

def tts(message):
    """
    This function takes a message as an argument and converts it to speech
    depending on the OS.
    """
    if sys.platform == 'darwin':
        tts_engine = 'say'
        return os.system(tts_engine + ' ' + message)
    elif sys.platform == 'linux2' or sys.platform == 'linux':
        tts_engine = 'espeak'
        return os.system(tts_engine + ' "' + message + '"')
```



## Rogelio Ferreira Escutia

Profesor / Investigador  
Tecnológico Nacional de México  
Campus Morelia



[rogelio.fe@morelia.tecnm.mx](mailto:rogelio.fe@morelia.tecnm.mx)



[rogeplus@gmail.com](mailto:rogeplus@gmail.com)



[xumarhu.net](http://xumarhu.net)



[@rogeplus](https://twitter.com/rogeplus)



[https://www.youtube.com/  
channel/UC0on88n3LwTKxJb8T09sGjg](https://www.youtube.com/channel/UC0on88n3LwTKxJb8T09sGjg)



[rogelioferreiraescutia](https://www.linkedin.com/in/rogelioferreiraescutia)

