Artificial Intelligent based Desktop Voice Assistant

UNIVERSITY OF ENGINEERING
&
MANAGEMENT, JAIPUR

Artificial Intelligent based Desktop Voice Assistant

Submitted in the partial fulfillment of the degree of

BACHELOR OF TECHNOLOGY

In

COMPUTER SCIENCE & ENGINEERING

Under

UNIVERSITY OF ENGINEERING & MANAGEMENT, JAIPUR

BY

ANKAN BISWAS

University Roll no: 12022002001040

University Registration no: 204202200200048

&

SANKALP SINGH

University Roll no: 12022002001138

University Registration no: 2042022002001138

&

AYUSH SAHA

University Roll no: 12022002001059

University Registration no: 204202200200064

UNDER THE GUIDANCE OF

PROF. SANTANU BASAK

COMPUTER SCIENCE & ENGINEERING

UNIVERSITY OF ENGINEERING & MANAGEMENT, JAIPUR

Approval Certificate

This is to certify that the project report entitled "AI based Desktop Voice Assistant" submitted by Ankan Biswas(Roll:12022002001040) ,Sankalp Singh(Roll:12022002001138) and Ayush Saha(Roll:12022002001059) in partial fulfillment of the requirements of the degree of **Bachelor of Technology** in **Computer Science & Engineering** from University **of Engineering and Management, Jaipur** was carried out in a systematic and procedural manner to the best of our knowledge. It is a bona fide work of the candidates and was carried out under our supervision and guidance during the academic session of 2022-2026.

_____

Prof. Santanu Basak
Project Guide, Assistant Professor (CSE)
UEM, JAIPUR

_____

Prof. Mrinal Kanti Sarkar
HOD (CSE)
UEM, JAIPUR

_____

Prof. A Mukherjee
Dean
UEM, JAIPUR

ACKNOWLEDGEMENT

The endless thanks go to Lord Almighty for all the blessings he has showered onto me, which has enabled me to write this last note in my research work. During the period of my research, as in the rest of my life, I have been blessed by Almighty with some extraordinary people who have spun a web of support around me. Words can never be enough in expressing how grateful I am to those incredible people in my life who made this thesis possible. I would like an attempt to thank them for making my time during my research in the Institute a period I will treasure. I am deeply indebted to my research supervisor, Professor Santanu Basak for encouraging me to work on such an interesting topic. Each meeting with him added in valuable aspects to the implementation and broadened my perspective. He has guided me with his invaluable suggestions, lightened up the way in my darkest times and encouraged me a lot in the academic life.

Ankan Biswas

Sankalp Singh

Ayush Saha

ABSTRACT

Communication and Technology has a long history but still constantly and actively growing and changing. The technology changes so fast that today everybody has an AI Personal assistant.[1]. The main goal of Artificial intelligence (AI) is the realization of natural dialogue between humans and machines. There are many IT companies have used the dialogue systems technology to set up various kinds of Virtual Personal Assistants (VPAs) based on their applications and areas for increasing interaction between human and machine, such as Microsoft's Cortana, Apple's Siri, Amazon Alexa, Google Assistant. We use python as a programming language because it has a major library which help executing commands. These AI assistant helps the user to do various task with ease, for example here in our project the user can open an installed appor search something on the internet or directly open a website and many more just with some simple voice command.

Table of Contents

CHAPTER - I

INTRODUCTION

Desktop voice assistants are software used to do certain tasks for the user, improve user experience, increase user productivity, etc. In short it is a software made for its user's ease. Typically, an Intelligent Personal Assistants will answer queries and perform actions via voice commands using a natural language user interface [1]. Virtual or Desktop assistants are in trend today because everyone wants that their work should be done efficiently with minimum stress and should have excellent quality. So, the virtual assistant is gaining an exponential popularity.

This basic idea behind creating this very Virtual assistant is to create standalone program that helps to the user to perform various task like

- Knowing what is the date and time
- Getting some specific information from Wikipedia
- Searching about something in the web browser
- Opening a website directly from the web browser and even playing some video from the YouTube website.
- Opening a preinstalled application and closing an opened application
- Terminating its own program.

This desktop voice assistant can do all of this by interacting with the user just using the user's voice and can also reply in voice.

CHAPTER - II

LITERATURE REVIEW

Bassam A, Raja N. et al, wrote about a very important statement and speech. The communication between human-machine programming is done with an analog signal that is converted into a speech signal into a digital wave. This technology is widely used, has unlimited usage and allows machines to respond appropriately and consistently to users' voices, and also provide useful and enjoyable services. Speech Recognition System (SRS) is slowly evolving and has endless applications. Research has revealed a summary of the process; is a simple model [1].

B. S. Atal and L. R. Rabiner et al, explained in terms of speech analysis, and the result is always supplemented with tone analysis. The study explained how to determine the pattern of determining whether a given piece of speech signal should be categorized as pronounced speech, non-verbal expression, or silence, depending on the size of the signal. The main limitation of the strategy is the need to use an algorithm in a set of selected sizes, as well as specific recording circulcumstance[2].

According to Deepak Shende, Ria Umabiya, the AIVA (Microsoft, Google Assistant from Google, and the recently appeared intelligent assistant under the name "AIVA" 2018) aimed at developing a voice-controlled personal assistant which is doing many things such as to search the Internet. It has some new features like posting comments on the social media websites such as Facebook, Twitter, etc. By just few simple commands. We can also know the weather around us and can get the climate conditions in your region [3].

Dhanraj, Vishal Kumar, and Semal Mahajan at 2022 et al. proposed a voice assistant using python speech to text module and had performed some api calls and system calls which has led to developing a voice assistant using python which allows the user to run any type of command through voice without interaction of keyboard. This can also run on hybrid platforms[4].

Reddy, Vyshnavi, Kumar and Soumya, march 2020. This study demonstrated that, An Intelligent Virtual Assistant (IVA) or Intelligent Personal Assistant (IPA) may software agent will perform the task or services on an individual. User may ask their assistants

questions, control home automation devices and media playbacks via voice and manage other basic task like emails, to do lists and calendar, with verbal commands[5].

Dhanraj, Kriplani and Mahajan, February 2022. This study examine, this voice assistant, their main problem and limitations. Also it describe that the method of creating a voice assistant without cloud service, which will show the expansion of such devices in the future. For the India in May 2022. The method was using such as speech recognition, API calls, Python backend. The finding shows that Virtual Assistant is the flexibility to contract for just the services they need, as like Alexa, Cortana, Siri, Google Assistant we also make virtual Assistant using Python for all window version[6].

CHAPTER - II

METHODOLOGY

II.I. System Architecture

The overall system design phase holds three phases:

- Voice analysis and conversion of the same to text.
- Data storage and processing
- Generating speech from the processed text output

II.II. Modulesused in the program are as follows:

- pyttsx3: is a text-to-speech conversion library in Python. Unlike alternative libraries, it works offline, and is compatible with both Python 2 and 3.[5]
- speech_recognition: Library for performing speech recognition, with support for several engines and APIs, online and offline.[6]
- datetime: The datetime module supplies classes for manipulating dates and times. While date and time arithmetic are supported, the focus of the implementation is on efficient attribute extraction for output formatting and manipulation.[7]
- wikipedia:Wikipedia is a Python library that makes it easy to access and parse data from Wikipedia.Search Wikipedia, get article summaries, get data like links and images from a page, and more. Wikipedia wraps the MediaWiki API so you can focus on using Wikipedia data, not getting it.[8]
- webbrowser: The webbrowser module supplies a high-level interface to allow displaying web-based documents to users.[9]
- AppOpener: AppOpener is the python library to open/close any application without knowing it's absolute path. The library works by making use of App Name and App Id. [10]

- googlesearch: googlesearch is a Python library for searching Google, easily. googlesearch uses requests and BeautifulSoup4 to scrape Google.[11]
- pyjokes:One-line jokes for programmers (jokes as a service) [12]

```
import pyttsx3
import speech_recognition as sr
import datetime
from datetime import date
import wikipedia
import webbrowser
from AppOpener import open, close
from googlesearch import search
import pyjokes
```

2.3 Functions that are included in the program are:

- **def speak(audio) –** This function is defined take the help of the pyttsx3 module to create ability of the model to convert text to speech.

```
def speak(audio):
    engine.say(audio)
    engine.runAndWait()
```

- **def whishMe() –** This function is defined to take the help of the datetime module to know what the time is and then the function speak() is used greet the user according to what time it is. Finally the function will end with another speak () giving its introduction.

```
def wishMe():
    hour = int(datetime.datetime.now().hour)
    if hour >= 0 and hour <12 :
        speak("Good Morning!")
    elif hour == 12  :
        speak("Good Noon!")
    elif hour >= 12 and hour <17 :
        speak("Good Afternoon!")
    else:
        speak("Good Evening!")

    speak("I am Veronica, your Little assistant. Please tell me how can i help you. ")
```

- def takeCommand() – This function is defined to take voice input from the user and convert it into text. This function uses the speech_recognition module's recognizer with the help of microphone to take the voice input. The audio is then converted to text using the speech_recognition module. This function returns the text as variable named query.

```
def takeCommand():
    r = sr.Recognizer()
    with sr.Microphone() as source:
        print("Listening...")
        r.pause_threshold=1
        audio = r.listen(source)

    try:
        print("Recognizing...")
        query = r.recognize_google(audio, language='en-in')
        print(f"User Said : {query}\n")

    except Exception as e:
        print("Say that again please...")
        return "none"

    return query
```

2.4 The main part of the program is as follows:

- The main part defines what program will do when specific command is encountered.

- The main starts with calling whishMe() to greet the user then it starts a infinite while loop
- Inside the loop, first takeCommand() is initiated. The returned text is stores in a variable named query and is converted to lower case.
- Then specific keyword is searched in the query, if found it the associated command is executed.

```
if __name__ == "__main__":

    wishMe()#calling fincton to wish the user

    while True:

        query = takeCommand().lower()#calling takeCommand can converting the returned


        if 'how are you' in query: #to answer and greet
            print("I am fine, Thank you")
            speak("I am fine, Thank you")
            print("How are you, Sir")
            speak("How are you, Sir")

        elif 'fine' in query or "good" in query:# to respond to the user
            speak("It's good to know that your fine")

        elif 'about your creator' in query:# to tell about its creator
            speak('Hey Veronica here, I am made by Aritra Sanyal and Ankan Biswas fres
            speak('University of Engineering & Management, Jaipur .')
```

```
elif 'date' in query: # to know what is the date
    today = date.today()
    date = today.strftime("%B %d, %Y")
    print(f"The date is {date}")
    speak(f"The date is {date}")

elif 'tell me a joke' in query: # to tell a joke
    My_joke = pyjokes.get_joke(language="en", category="neutral")
    print(My_joke)
    speak(My_joke)

elif 'wikipedia' in query: #logic or searching wikipedia
    speak('Searching Wikipedia...')
    query = query.replace("wikipedia", "")
    results = wikipedia.summary(query, sentences=3)
    speak("According to Wikipidia")
    print(results)
    speak(results)
```

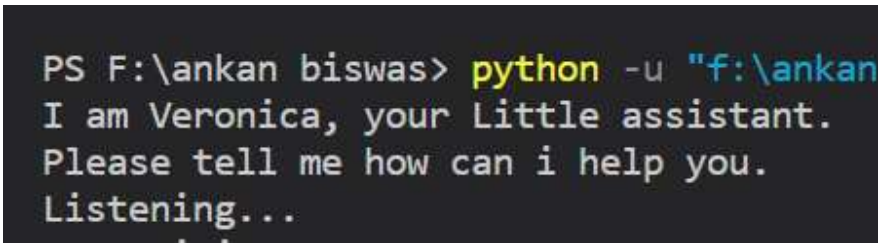
```
elif 'search' in query:# to search something in browser
    webbrowser.open(query.replace('search', ''))

elif 'open' and 'website' in query:# to open a specific website
    for link in search(query,tld='com',num=1,stop=1):
        pass
    webbrowser.open(link)

elif 'open' in query: # to  open a preinstalled app
    open(query.replace('open',''))

elif 'close' in query: # to close a open app
    close(query.replace('close',''))

elif 'exit'in query or 'bye' in query: #to exit the program
    print("Thanks for giving me your time")
    speak("Thanks for giving me your time")
    exit()
```

CHAPTER -III

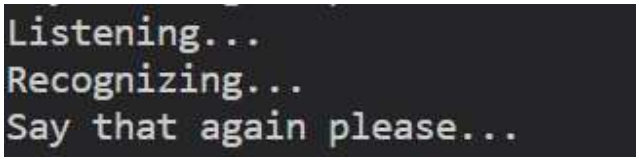
RESULT

3.1. Working Principle of the proposed model

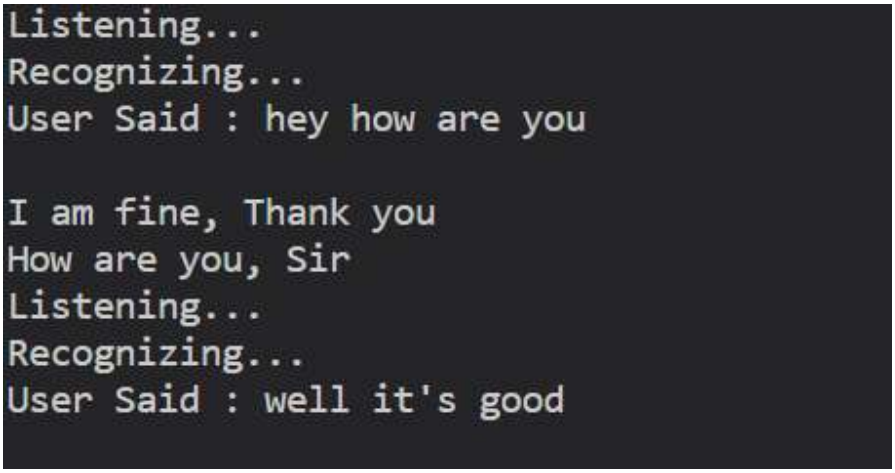
- Upon running the program, it first greets us and start to listen to our command



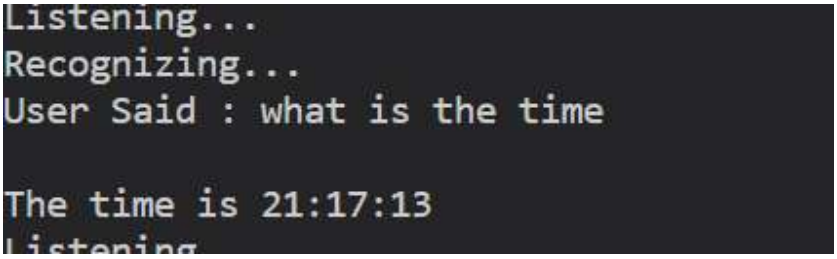
- The program will try to listen to what the user says anything or not while pausing for a bit if the command is not listed or isn't understandable, the program displays "Say that again please..." and again tries to listen what the user is trying to say



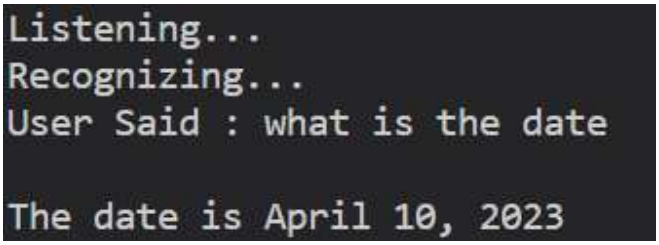
- If the keyword "how are you" is found in the query or the input speech it responds with



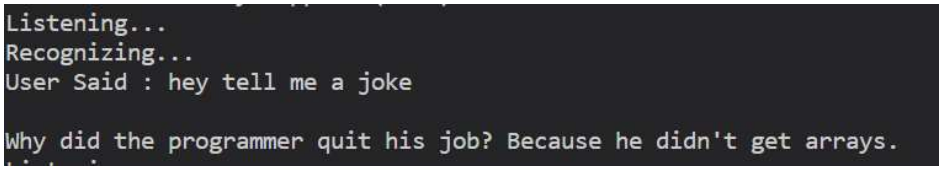
- If the keyword "the time" isfound in the query or the input speech it responds with



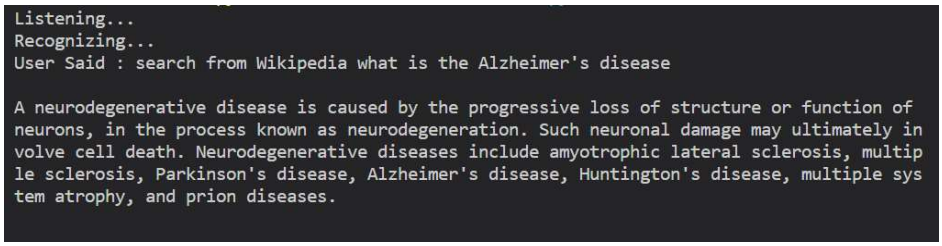
- If the keyword “date” is found in the query or the input speech it responds with



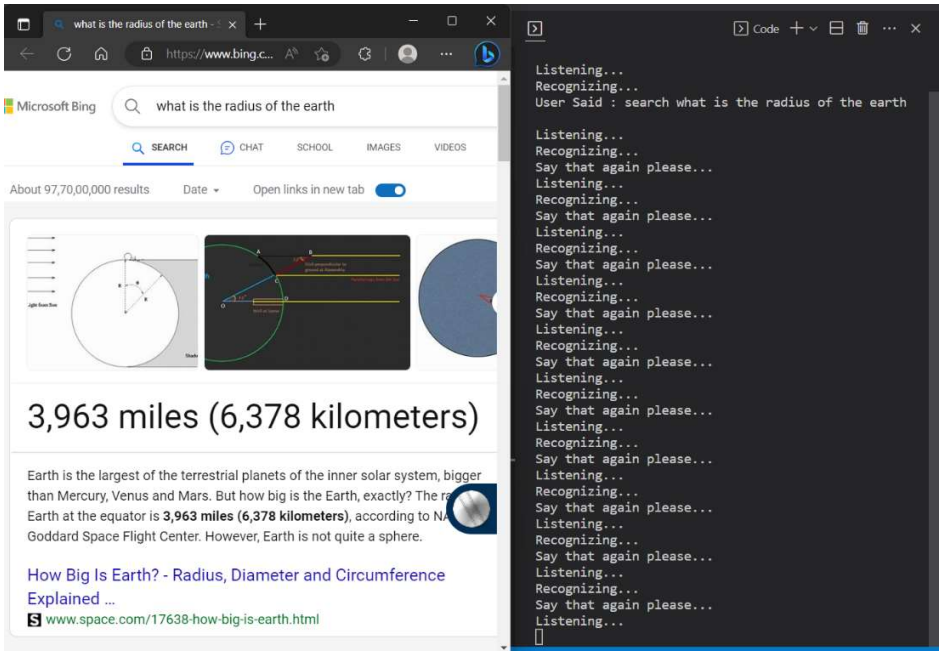
- If the keyword “joke” is found in the query or the input speech it responds with



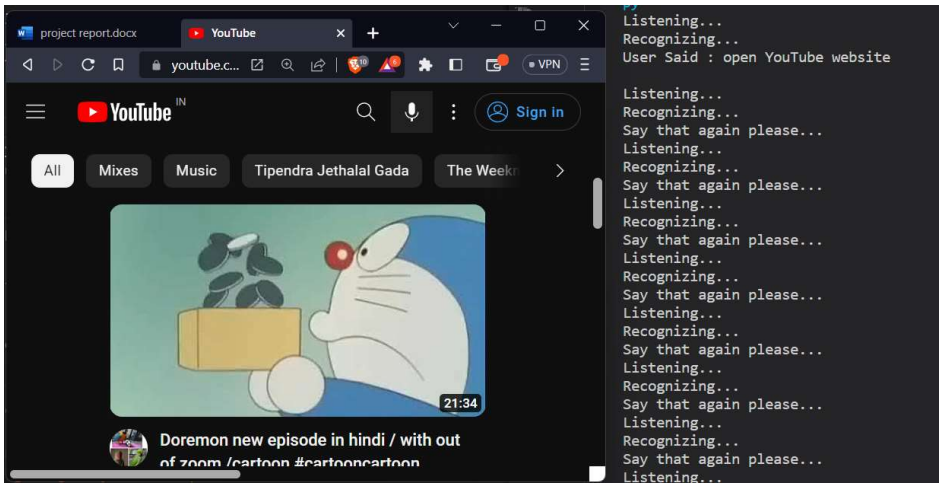
- If the keyword “Wikipedia” is found in the query or the input speech it searches the query using the Wikipedia module and responds with



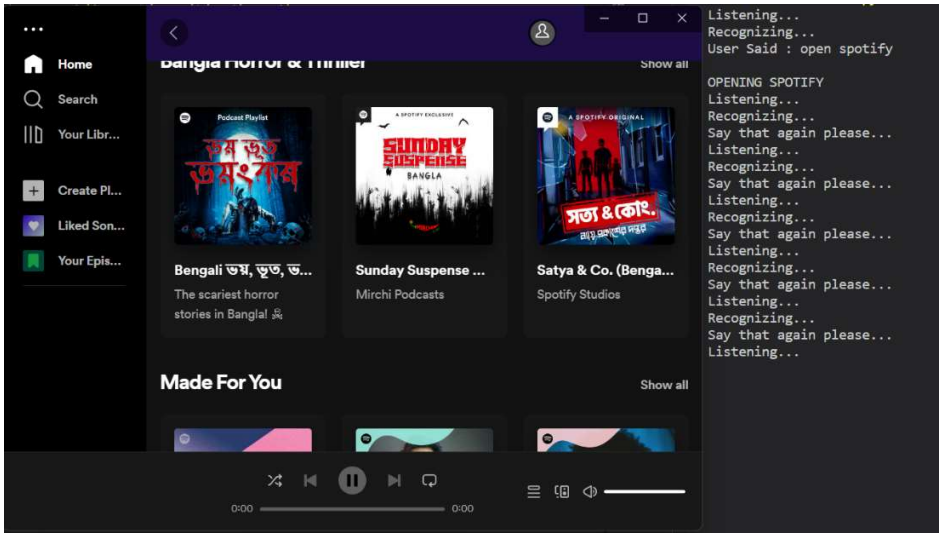
- If the keyword “search” is found in the query or the input speech it opens the web browser module and searches the query.



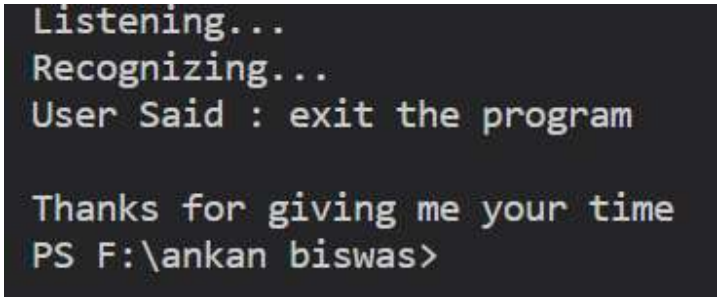
- If the keywords “open” and “website” is found in the query or the input speech it opens the web browser and opens the specific website



- If the keyword “open” is found in the query or in the input speech, it opens the specific preinstalled app if the app is not present then the app is shows that the app is not found.



- If the keyword “close” is found in the query or in the input speech then the opened app closed and if the app or if the app is not open it displays that the app is not open.
- If the keyword “exit” or “bye” is found in the query or in the input speech then the program terminatesitself.



CONCLUSION& FUTURE SCOPE

In this report we have discussed about AI based desktop voice assistant using python. It performs basic tasks like telling date, time, simple interacting with the user, search Wikipedia, searching on the web browser,opening a specific website or opening or closing a specific app.

The current proposed model has few limitations that can be considered in the later revised model that include:-

- a graphical user interface,
- the flexibility to type instead of speaking,
- operating offline
- Operating commands inside a preinstalled application,
- incorporation with machine learning, neural network and IOT
- Incorporation with Bing Chat or Google Bard or being itself such that it becomes a software like Bing Chat or Google Bard but uses voice input and output and just like Alexa, Siri or Google Assistant be on every single phone or any device imageable.

REFERENCE

[1] Webopedia, web page [Online]. Available: https://www.webopedia.com/TERM/I/intelligent-personal-assistant.html

[2] Bansal, Tushar, et al. "GENESIS-THE DIGITAL ASSISTANT (PYTHON)."

[3] P.Praddeep, P.Balaji, S.Bhanumathi, “International Journal of Recent Technology and Engineering”, ISSN: 2277-3878, September 2019

[4] Dhanraj, Vishal Kumar, and Semal Mahajan. "Research Paper onDesktop Voice Assistant." (2022).

[5] https://pypi.org/project/pyttsx3/

[6] https://pypi.org/project/SpeechRecognition/

[7] https://docs.python.org/3/library/datetime.html

[8] https://pypi.org/project/wikipedia/

[9] https://docs.python.org/3/library/webbrowser.html?highlight=webbrowser#module-webbrowser

[10] https://pypi.org/project/appopener/

[11] https://pypi.org/project/googlesearch-python/

[12] https://pypi.org/project/pyjokes/