

Plagiarism Scan Report

Report Generated on: Jul 18,2022



Total Words:	911
Total Characters:	5535
Plagiarized Sentences:	3.57
Unique Sentences:	47.43 (93%)

Content Checked for Plagiarism

Wikipedia Article Extraction Using Voice Commands

Abstract :-

This project uses python as primary language. This project will aim to create an application which will enable person using it to extract information from Wikipedia, and the application will read important points from it to the user and the whole article will be displayed. Moreover this project can be expanded into many sectors and it can be more than just a Wikipedia article extractor.

Introduction:-

This project is made using the python language which is best for web scraping projects and to reduce the complexity of the code. This project will make the process of searching an article from Wikipedia easier by giving you the option to extract the article from Wikipedia by just a voice command.

Different Creators have made different modules to help the programmer to create projects with ease. These modules are used in this project to make it with ease, these modules were downloaded externally using the PIP command provided by python developers.

Different functions are created in this project and each function performs different operations, these operations are created for the smooth functioning of the project.

How Wikipedia works and What is web scraping:-

1. Wikipedia Working :- Wikipedia works as an encyclopaedia which is widely accessible to people who have access to internet. Wikipedia is an online platform which provides different articles on various topics. These articles are written and edited collaboratively. Anonymous volunteers from all around the world help write and edit the Wikipedia articles. There are some sections of Wikipedia which are restricted to edit to prevent disruption or vandalism. Anyone is allowed to add or edit words, references, images, and other media here. The content must not be copyright restrictions and contentious material about living people. It must conform with Wikipedia's policies, including being verifiable against a published reliable source. If someone damages the contents of the Wikipedia, then their software is built to change it easily. The changing of the article is just one EDIT button away.
2. Web Scraping – Assume you read some information or article on internet, and you want to use the article or the information provided on the website in your project. That is when web scraping comes into play, it is a technique to extract the data from a particular website. This data can be in any form text, tabular format, numerical data, etc. This can also extract any particular form of data from a website. This technique is basically used to extract large amounts of data from a website. It has many uses for example:- Train a Machine Learning Algorithm, Stock Market Analysis, etc. Web scraping is an automatic system to gain large quantities of data from websites. utmost of this data is unshaped data in an HTML format which is also converted into structured data in a spreadsheet or a database so that it can be used in colourful operations. There are numerous different ways to perform web scraping to gain data from websites. These include using online services, particular API’s or indeed creating your law for web scraping from scrape. numerous large websites, like Google, Twitter, etc.

The Modules and how they work:-

Before getting to know all the modules we must know which Language will I be using. I chose Python as a programming language as web scraping is easy and it provides various modules to make my code easy.

So the list of all the modules required and used in the project are:-

1) Datetime Module:- In Python, date and time are not a data type of their own, but a module named datetime can be imported to work with the date as well as time. Python Datetime module comes built into Python, so there is no need to install it externally. It contains 6 main categories :-

1. Date – This displays current Date, it follows Gregorian calendar
2. Time – This displays current Time according to the 24hr clock format.
3. Datetime – This displays the current date and time.
4. Timedelta - A duration expressing the difference between two date, time, or datetime instances to microsecond resolution.

2) Speech Recognition:- Speech Recognition is an important feature in several applications used such as home automation, artificial intelligence, etc. This article aims to provide an introduction on how to make use of the Speech Recognition library of Python. This is useful as it can be used on microcontrollers such as Raspberry Pi's with the help of an external microphone.

3) Pyttsx3:- It is a module which is used to convert the speech to text. This library can also work offline while other libraries require the user to be online. For every operating system there are different voices. Like Windows- "SAPI5", Mac- "nsss", Linux-"espeak".

4) Wikipedia Module:- In this article, we will see how to use Python's Wikipedia module to fetch a variety of information from the Wikipedia website. This basically uses web scraping to extract the article from the website of Wikipedia. This returns the result in the format of a string. Ex:-

Python code of the project:-

Output of the Code:-

1.

2.

3.

Conclusion:-

Conclusion of this project is that I learned about how advance Python language can be. A project to extract Wikipedia by voice commands was created, and it contains very minimum errors. I used different examples and phrases to check the program if it is working properly. At the

end this was a successful project.

About Wikipedia

Anyone is allowed to add or edit words, references, images, and other media here.

100%

https://en.wikipedia.org/wiki/Wikipedia:About

About Wikipedia

It must conform with Wikipedia's policies, including being verifiable against a published reliable source.

100%

https://en.wikipedia.org/wiki/Wikipedia:About

www.geeksforgeeks.org › python-datetime-modulePython datetime module - GeeksforGeeks

Nov 17, 2021 · Python datetime module. In Python, date and time are not a data type of their own, but a module named datetime can be imported to work with the date as well as time. Python Datetime module comes built into Python, so there is no need to install it externally. Python Datetime module supplies classes to work with date and time.

42%

https://www.geeksforgeeks.org/python-datetime-module

github.com › Rawat30 › speech-recognitionGitHub - Rawat30/speech-recognition: Speech Recognition is ...

Oct 30, 2021 · Speech Recognition is an important feature in several applications used such as home automation, artificial intelligence, etc. This article aims to provide an introduction on how to make use of the SpeechRecognition library of Python.

59%

https://github.com/Rawat30/speech-recognition