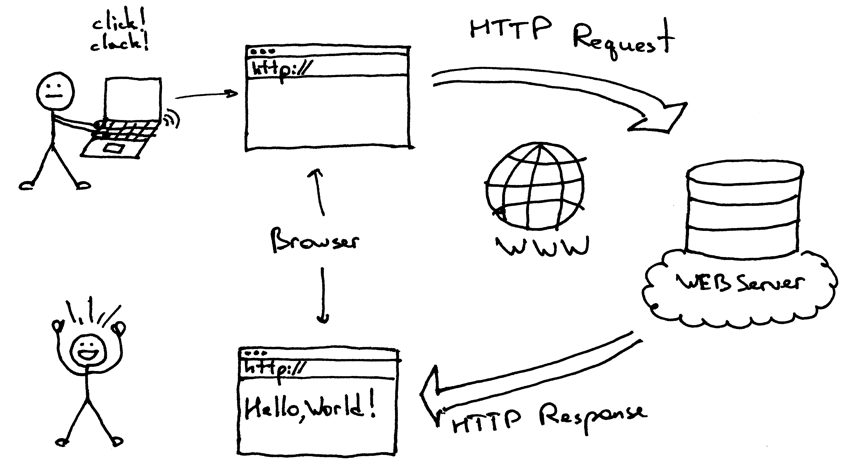
Each step to create a program that fetches repositories from a GitHub user, scans the files for specific patterns, and presents the findings. By the end, you will have a fully functional tool that can find URLs, emails, phone numbers, API keys, usernames, and passwords within GitHub repositories.  
**1.Introduction:**

  
Web scraping is a powerful technique for extracting information from websites. In this tutorial, we will use Python to build a tool that scrapes repositories from a GitHub user, scans files for specific patterns, and displays the results. This project is an excellent way to learn about APIs, regular expressions, and data processing.

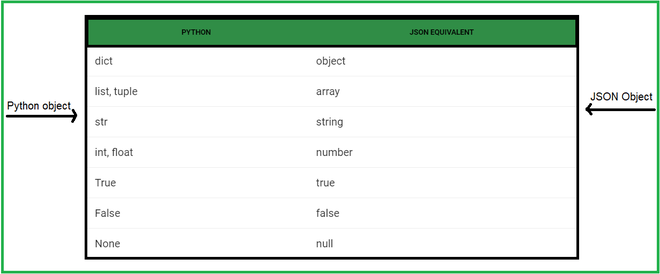
**2. Setting Up the Environment:**First, ensure you have Python installed on your system. You will also need the following libraries:

* requests: For making HTTP requests to the GitHub API
* re: For regular expression operations
* json: For handling JSON data
* time: For tracking execution time
* tqdm: For displaying progress bars (library).

Install the required libraries using pip:

pip install requests tqdm

# Note : what is Json in python ? Python JSON

**Python JSON JavaScript Object Notation** is a format for structuring data. It is mainly used for storing and transferring data between the browser and the server. [Python](https://www.geeksforgeeks.org/python-programming-language/learn-python-tutorial/) too supports JSON with a built-in package called JSON. This package provides all the necessary tools for working with JSON Objects including parsing, [serializing](https://www.geeksforgeeks.org/serializing-json-data-in-python/)(Serialization is the process of encoding the from naive data type to JSON format), deserializing(vise versa), and many more.   


**Let’s take a look at how we serialize Python data to JSON format with these methods:**

* Dump().
* Dumps().

## json.dump()

[**json.dump()**](https://www.geeksforgeeks.org/json-dump-in-python/) method can be used for writing to JSON file. Write data to a file-like object in json format.

***Syntax:****json.dump(dict, file\_pointer)*

***Parameters:***

* ***dictionary –****name of dictionary which should be converted to JSON object.*
* ***file pointer –****pointer of the file opened in write or append mode.*

**Below is the implementation:**

Converting python object and writing into json file

|  |
| --- |
| # import module  import json    # Data to be written  data = {      "user": {          "name": "satyam kumar",          "age": 21,          "Place": "Patna",          "Blood group": "O+"      }  }    # Serializing json and  # Writing json file  with open( "datafile.json" , "w" ) as write:      json.dump( data , write ) |

**Output:**

https://media.geeksforgeeks.org/wp-content/uploads/20201125190923/Capture1.PNG

*data\_file.json*

## json.dumps()

[**json.dumps()**](https://www.geeksforgeeks.org/json-dumps-in-python/) method can convert a Python object into a JSON string.

***Syntax:****json.dumps(dict)*

***Parameters:***

* ***dictionary –****name of dictionary which should be converted to JSON object.*

**Below is the implementation:**

Converting python object into json string.

|  |
| --- |
| # import module  import json    # Data to be written  data = {      "user": {          "name": "satyam kumar",          "age": 21,          "Place": "Patna",          "Blood group": "O+"      }  }    # Serializing json  res = json.dumps( data )  print( res ) |

**Output:**

# https://media.geeksforgeeks.org/wp-content/uploads/20201125191430/Capture-660x37.PNG

# 3. Fetching GitHub Repositories: Serializing json