ASSIGNMENT

PROG 301: Advanced Programming with Python and Scripting

PREPARED BY

Mehrin Farzana
ID: 2101013
Session: 2021-2022
Date: 11/12/2024

SUPERVISED BY

Farzana Akter
Assistant Professor
Department of IRE, BDU

BANGABANDHU SHEIKH MUJIBUR RAHMAN DIGITAL UNIVERSITY

(BDU)

Code:

```
import urllib.request
import json
from bs4 import BeautifulSoup
import os

url = "http://example.com"
response = urllib.request.urlopen(url)
webContent = response.read()

soup = BeautifulSoup(webContent, 'html.parser')
data = {}

heding = soup.select_one('h1')
data['title'] = [p.get_text() for p in paragraphs]

paragraphs = soup.select_one('p')
data['paragraphs'] = [p.get_text() for p in paragraphs]

paragraphs = soup.select_one('a')
data['link'] = [p.get_text() for p in paragraphs]

json_data = json.dumps(data, indent=4)

with open('output.json', 'w') as json_file:
    json_file.write(json_data)
file_path = os.path.abspath('output.json')
print(f"Data saved to {json_file}")
print(f"File is located at: {file_path}")
print("Data has been written to output.json")
with open('output.json', 'r') as json_file:
    print(json_file.read())
```

Output:

```
Data saved to <_io.TextIOWrapper name='output.json' mode='w' encoding='UTF-8'>
File is located at: /content/output.json
Data has been written to output.json
{
    "title": [
        "More information..."
    ],
    "paragraphs": [
        [
            "This domain is for use in illustrative examples in documents. You may use this",
            "    domain in literature without prior coordination or asking for permission."
        ]
    ],
    "link": [
        "More information..."
    ]
}
```