jupyter Google image Scrapping Last Checkpoint: 1 minute ago

File Edit View Run Kernel Settings Help

Trusted

Code

JupyterLab Python 3 (ipykernel)

[ ]:

```
from flask import Flask, render_template, request,jsonify
from flask_cors import CORS,cross_origin
import requests
from bs4 import BeautifulSoup
from urllib.request import urlopen as uReq
import logging
import pymongo
logging.basicConfig(filename="scrapper.log" , level=logging.INFO)
import os

app = Flask(__name__)

@app.route("/", methods = ['GET'])
def homepage():
    return render_template("index.html")

@app.route("/review" , methods = ['POST' , 'GET'])
def index():
    if request.method == 'POST':
                try:

                    # query to search for images
                    query = request.form['content'].replace(" ","")

                            # directory to store downloaded images
                    save_directory = "images/"

                            # create the directory if it doesn't exist
                    if not os.path.exists(save_directory):
```

jupyter Google image Scrapping Last Checkpoint: 1 minute ago

File Edit View Run Kernel Settings Help

Trusted

Code

JupyterLab Python 3 (ipykernel)

```
        # directory to store downloaded images
save_directory = "images/"

        # create the directory if it doesn't exist
if not os.path.exists(save_directory):
    os.makedirs(save_directory)



        # fake user agent to avoid getting blocked by Google
headers = {"User-Agent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/58.0.3029.110 Safari

        # fetch the search results page
response = requests.get(f"https://www.google.com/search?q={query}&sxsrf=AJOqlzUuff1RXi2mm8I_OqOwT9VjfIDL7w:1676996143273&source=lnms&t


        # parse the HTML using BeautifulSoup
soup = BeautifulSoup(response.content, "html.parser")

        # find all img tags
image_tags = soup.find_all("img")

        # download each image and save it to the specified directory
del image_tags[0]
img_data=[]
for index,image_tag in enumerate(image_tags):
            # get the image source URL
            image_url = image_tag['src']
            #print(image_url)

            # send a request to the image URL and save the image
```

jupyter Google image Scrapping Last Checkpoint: 1 minute ago

File Edit View Run Kernel Settings Help

Trusted

Code

JupyterLab Python 3 (ipykernel)

```
                for index,image_tag in enumerate(image_tags):
                        # get the image source URL
                        image_url = image_tag['src']
                        #print(image_url)

                        # send a request to the image URL and save the image
                        image_data = requests.get(image_url).content
                        mydict={"Index":index,"Image":image_data}
                        img_data.append(mydict)
                        with open(os.path.join(save_directory, f"{query}_{image_tags.index(image_tag)}.jpg"), "wb") as f:
                            f.write(image_data)


                return "image laoded"
            except Exception as e:
                logging.info(e)
                return 'something is wrong'
        # return render_template('results.html')

    else:
        return render_template('index.html')


if __name__ == "__main__":
    app.run(debug=True)
```