CYCOGNITO

Web Browser Module Home Task

You are requested to create a docker image called “browser_module” which programmatically browses to a list of urls provided in an input file and implements a certain functionality which will be explained below.

Running the docker image should look like this:

```
docker run -v /path/to/input/dir:/input -v /path/to/output/dir:/output -it hometask/browser_module
```

You can assume that within the docker container under the path “/input” there will be a file called “urls.input” containing a list of urls for the program to scan one in each line(Can be any amount of urls). Example for such file:

```
http://www.bbc.com/
http://www.cnn.com/
http://www.cycognito.com/
http://www.google.com/
http://www.yahoo.com/
```

The program will create a directory for each url (“url_1”, “url_2”, ...) under the path “/output” in the container’s filesystem which will be the output directory of the relevant url.

Functionality:

Part 1:
HTML - The html content of the web page, return it as a string within a JSON file called “browse.json” as “html”. Save the file inside the relevant output directory(“url_1”, “url_2”, ...)

Part 2:
Web Resources - A list of urls accessed during page loading for example images, links etc. return as a list within the same “browse.json” file as “resources” and for each resource its status code(like status=200).

Part 3:
Screenshot - A screenshot of the web page, placed as a “screenshot.png” file within the relevant directory.

Part 4:
Decoded screenshot - Write the above “screenshot.png” in the form of a decoded base64 string as “screenshot” inside “browse.json”

CYCOGNITO

Solving the task:
We wouldn't want you to put in too much of your time into this task so please return it within **3 days** from when it was sent to you. It is fine if not all of parts 1-4 were implemented during this time frame.

Requirement/ Guidelines:

- Use Python as the programming language
- Use Chrome as the web browser and make sure it's headless.
- Write tests.
- Put all the dependencies within the Dockerfile/requirements.txt so no installations will need to be done in order to run it.
- Attach a **README.md** file explaining how to build and run the Docker image.
- Send it all as a link of a private Github repository

Have Fun!