

TP final

Ressource R2.08

Durée : 60 minutes

Consignes : Durant la séance de TP, vous pourrez vous appuyer sur le cours, les TD et les TP et sur le site web du module re, dont l'adresse est fournie dans la deuxième question. Tout autre document est interdit. Deux fichiers python sont mis à votre disposition pour faire ce TP, le premier est `AnalyseurWeb.py` et qui contient la contenu de la classe du même nom et le fichier `main.py`. Vous pourrez tester dans votre programme les URL suivantes :

- `www.uha.fr`,
- `www.scientist.fr`,
- `www.irimas.uha.fr`,
- `www.labri.fr`.

1 Extraction d'information d'un site web

1. (2 points) Compléter la fonction `__check_machine_name()` qui vérifie si l'url passée en paramètre respecte le schéma suivant : `www.domaine.tld`, avec `tld` qui peut être une chaîne qui contient exclusivement deux ou trois caractères. Par exemple `www.uha.fr` est une URL valide.
2. (1 point) À partir de la documentation https://developer.mozilla.org/fr/docs/Learn/Common_questions/What_is_a_URL, compléter la fonction `open_url(self, machine_name) -> str:` afin de pouvoir retourner la page web demandée.
3. (1 point) Écrire la fonction `nbr_css() -> bool` qui compte le nombre de fichiers css chargés par la page web donnée en paramètre. Vous pourrez utiliser la touche F12 de Firefox pour vous aider à trouver les balises ainsi que le <https://docs.python.org/3/library/re.html#functions> pour les fonctions du module `re`.
4. (1 point) Écrire la fonction `get_langage() -> str` qui retourne la langue utilisée pour le site web en entier.

2 Historique de navigation

1. (1 point) Mettre en place un système qui permet de mémoriser le nombre de fois où le site a été visité avec la fonction `open_url()`.
2. (2 points) Créer une fonction qui retourne les trois sites les plus visités.
3. (2 points) Créer une fonction qui sauvegarde l'historique des sites visités.