

Comment scrapper avec Screaming Frog ?

Screaming frog peut aussi scrapper du contenu. ça veut dire qu'il peut copier une portion précise de la page. Il est par exemple possible de récupérer les adresses sur les pages de 250 vendeurs de [letzshop.lu](https://www.letzshop.lu) et les lister dans un tableur.

En pièce jointe, le fichier de configuration pour l'extraction des données Letzshop en vue d'automatiser la création d'ads Adwords

- Ouvrir SF
- Limiter le crawl aux pages qui contiennent l'information via une expression régulière. Par exemple https://letzshop.lu/fr/vendors.* Plus d'info sur la doc SF - <https://www.screamingfrog.co.uk/seo-spider/user-guide/configuration/#include> - <https://youtu.be/7G-Ds8vhUvE?t=22>
- Via les developper tool, isoler le contenu à scrapper (CSSpath le plus fréquemment). Plus d'information sur la doc SF pour effectuer les custom extractions - <https://www.screamingfrog.co.uk/seo-spider/user-guide/configuration/#custom-extraction> - <https://youtu.be/7G-Ds8vhUvE?t=63>
- Lancer le crawl - <https://youtu.be/7G-Ds8vhUvE?t=144>
- Accéder à l'extraction custom via l'onglet custom > extraction et exporter en CSV - <https://youtu.be/7G-Ds8vhUvE?t=209>

Fichier check des commerçants:

https://docs.google.com/spreadsheets/d/1jL8kCLAm4wcp1I3XFhGNLBi29N-fIC-QWNZ-f_98aXE/edit#gid=144726851

Révision #3

Créé 1 September 2020 13:10:49 par Patrice Viale

Mis à jour 27 May 2022 13:19:59 par Jean-Paul Diaz-Caneja