



FORMATIONS WEB SCRAPING

Apprenez à récupérer de la donnée en masse sur le web, que vous soyez développeur ou non

2019 - 2020

scrap-and-drop.fr





L'accès à la donnée est un sujet de plus en plus central. A ce titre, le **web scraping** représente la solution à de nombreux enjeux.

Que ce soit pour récupérer de l'information, faire grossir vos bases de **prospects**, mieux connaître **le marché** sur lequel vous œuvrez, ou encore **automatiser** des tâches vous prenant trop de temps, l'acquisition de cette compétence vous permet de gagner en **autonomie** (et donc...de faire des **économies** ;)).

Nous proposons donc plusieurs types de **formations**, adaptées en fonction de votre profil.

Ces formations sont proposées en collaboration avec l'organisme "**Fixit Formations**". Que vous soyez **salarié**, **indépendant** ou en **recherche d'emploi**, **il y a de fortes chances que vous puissiez être financé**, et que vous n'ayez donc **rien à sortir de votre poche !**

Pour que l'on fasse le point ensemble sur votre cas précis, rendez-vous à la dernière page de ce catalogue, où vous trouverez nos coordonnées.





RECUPERER DE LA DONNEE SUR LE WEB SANS CODER

Public: Personnes n'ayant aucune compétences en programmation web

JOUR 1

- Rappel sur le concept et la pratique du web scraping
- Définition de quelques concepts clés : client/serveur, requêtes HTTP, code source, site dynamique...
- Qu'est ce qu'une API ?
- Présentation des 3 types de cibles que vous allez rencontrer sur le web
- Comment définir le "type" auquel appartient le site que je souhaite scraper ? Apprendre à se servir correctement de son navigateur web et de ses outils.

JOUR 2

- Type 1 : l'API. Comment tester une API avec Postman et récupérer directement de la donnée.
- Type 2 : Le site web statique. Récupérer de la donnée avec l'extension Web Scraper.
- Convertir les données récupérées aux formats texte, CSV, ou les stocker en base de données MySQL
- Type 3 : les sites web dynamiques : quels outils ?

INITIATION AU WEB SCRAPING EN PHP

Public: Personnes n'ayant aucune compétences en programmation web et souhaitant s'initier au PHP afin d'être en mesure de créer des scripts de scraping

JOUR 1

- Rappel sur le concept et la pratique du web scraping
 - Définition de quelques concepts clés : client/serveur, requêtes HTTP, code source, site dynamique...
 - Qu'est ce qu'une API ?
 - Présentation des 3 types de cibles que vous allez rencontrer sur le web
 - Comment définir le "type" auquel appartient le site que je souhaite scraper ?
- Apprendre à se servir correctement de son navigateur web et de ses outils.

JOUR 2

- Type 1 : l'API. Comment tester une API avec Postman.
- Présentation du langage PHP, introduction à la programmation : notions, logiques et concepts clés.
- Mettre en place un environnement de développement sur son ordinateur: installer un serveur local et un logiciel de base de données (WAMP/LAMP et MySQL Workbench)
- Type 1 : créer un script PHP simple pour récupérer de la donnée sur une API.

JOUR 3

- Type 2 : créer un script PHP pour récupérer de la donnée directement sur un site web statique en parcourant le code HTML.
- Type 3 : les sites web dynamiques : quelques pistes.
- Exporter la donnée récupérée aux formats texte et CSV.
- Savoir créer, configurer et alimenter une base de données MySQL via le logiciel Workbench.





FORMATION AVANCÉE AU WEB SCRAPING SOUS NODEJS

Public: Développeurs web maîtrisant un minimum le JavaScript

JOUR 1

- L'analyse préalable à tout processus de scraping : présentation de notre méthodologie.
- Apprendre à mieux connaître notre cible: analyse des requêtes, du code source, utilisation de Postman...
- Récupérer de la donnée brute sur une API avec la librairie Axios.
- Requêtes imbriquées: rappel au sujet de l'asynchrone.

JOUR 2

- Scraper un site web en parcourant son code HTML avec la librairie Cheerio.
- Export des données en MySQL et CSV.
- Scraper un site web dynamique avec un headless-browser : présentation de la librairie Puppeteer.

JOUR 3

- Scraper un site web dynamique avec Puppeteer (suite).
- Bonnes pratiques et optimisation des scripts de scraping : les temps morts.
- Bonnes pratiques et optimisation des scripts de scraping : le User-agent.
- Bonnes pratiques et optimisation des scripts de scraping : la rotation d'IPs.

La structure et le contenu des formations présentées dans ce catalogue peuvent être adaptés en fonction de vos besoins. N'hésitez pas à nous contacter pour plus d'informations.

Votre interlocuteur



Frédéric Ferri

Développeur web full-stack
Spécialiste en web scraping



contact@scrap-and-drop.fr



06.67.57.33.79



<https://scrap-and-drop.fr>

