

Scraping et extraction de données Web



STG-E18-013

Durée : 3 jours

Niveau : Intermédiaire

Prix : 1 770 €

OBJECTIFS

Le Scraping est le fait de récupérer de façon automatisée des données structurées ou semi-structurées depuis une page Web ou une interface. Cette technique est largement répandue dans les entreprises.

Les objectifs sont de maîtriser les bases du traitement de données afin de mettre en œuvre une technique d'acquisition de données automatisée.

SESSION(S) PLANNIFIEE(S)

- 10, 11 et 12 juillet 2018 – (*Bordeaux*)

[S'inscrire sur le site Internet de la formation](#) ou par téléphone : 07 64 14 76 12

Formation **disponible toute l'année** à la demande.

PREREQUIS

- Connaissances de base de la programmation Python
- Notions en HTML, CSS, XML, JSON, XPath et Regex

PROGRAMME DETAILLE

Introduction

- Législation
- OpenData et licences
- Propriété des données

API

- Principe de l'API
- Requêtes avec R et Python

Web

- Principe du scraping
- Structure des pages Web
- Exemples avec R et Python

Outils

- Scrapy
- PhantomJS
- BM

Avancé

- Proxy et authentification
- Gestion des erreurs
- Ajax

LOGICIELS UTILISES

Python, R, Google Chrome

Compatible PC & MAC
Logiciels non payants

MODALITES PEDAGOGIQUES

Afin de permettre aux étudiants d'acquérir des compétences opérationnelles en entreprise, cette formation associe des connaissances notionnelles (exposés théoriques) avec des savoir-faire pratiques (applications, cas concrets, travaux pratiques).

Vous souhaitez obtenir plus de détails sur cette formation ? Rendez-vous sur notre site Internet, [sur la page dédiée à la formation](#).

Pour toute question, n'hésitez pas à contacter conseil@datauniversity.fr et à consulter notre site Internet <https://datauniversity.fr>