



# Extraire les données en datajournalisme

Ref 4610219

## Extraire les données en datajournalisme

Formations courtes

---

Approfondissement

1 jour (7 heures)

Paris

---

**Récupérer des données disséminées dans un site, intégrées dans un document PDF ou une image scannée.**

## Objectifs

- Récupérer des données disséminées dans le cadre d'un projet en datajournalisme.
- Extraire des données quel que soit leur format.

## Pour qui ?

Journalistes tous médias. Tous producteurs de contenus digitaux travaillant dans une structure publique.

## Prérequis

Avoir suivi la formation "L'outil du datajournaliste: le tableur" ou maîtriser le tableur dans ses fonctions principales.

Code dokélio :



## Homologations

# Programme

## Extraire depuis un site web

Découvrir Outwit: se familiariser avec ce logiciel qui permet de récupérer des informations et des données de manière structurée puis les transformer et les stocker.

Fonctions d'import depuis Google spreadsheet.

Utiliser Kimono: se familiariser avec le data scraping.

## Extraire depuis un PDF

PDF to Excel, pdftables, etc.

Découvrir Tabula.

## Extraire depuis un scan, une image

S'initier aux logiciels OCR.

Extraire les données en appliquant les règles déontologiques.

## Nombreux exercices et cas pratiques

700€ HT

## Prochaines sessions

### Paris

07 nov. 2019

07-11-2019

07-11-2019

S'inscrire

[Nous contacter](#)