

---

---

# WEB SCRAPING : TABLEAU COMPARATIF ENTRE PYTHON, NODE ET R

VISION CROISÉE SUR LES 3 LANGAGES LES PLUS PLÉBISCITÉS DANS LE DOMAINE DU WEB SCRAPING

---

---

REPARTITION DES RESULTATS PAR LANGAGE POUR LA RECHERCHE "WEB SCRAPING" SUR YOUTUBE\*



\*Sur les 20 premiers résultats observables au 02/03/2020. Les résultats restants concernent majoritairement des vidéos utilisant des outils en no-code.

---

---

## PRINCIPALES LIBRAIRES

### Python

- Selenium
- Scrapy
- BeautifulSoup
- LXML
- Urllib

### Node

- Cheerio
- Puppeteer
- Nightmarejs
- Osmosis
- Request

### R

- RCrawler
- Rvest

LIBRAIRIES PERMETTANT DE FAIRE DU HEADLESS-BROWSER (CONTENU DYNAMIQUE)

### Python

- Selenium

### Node

- Puppeteer
- Nightmarejs

### R

- V8\*
  - RSelenium\*
- 
- 

## PYTHON

### POINTS FORTS



MULTITHREADING POSSIBLE



LANGAGE FLEXIBLE, PROPRE ET INTUITIF



LANGAGE PUISSANT POUR FAIRE DE L'ANALYSE DE DATAS



FORTE COMMUNAUTÉ ET RESSOURCES NOMBREUSES

## NODE / JAVASCRIPT

### POINTS FORTS



GRAND NOMBRE DE LIBRAIRIES, LANGAGE EN EXPANSION



VISION A 360, PERMET D'ETRE OPERATIONNEL SUR FRONT ET BACK



LANGAGE TRES PERFORMANT POUR FAIRE DU HEADLESS BROWSER



PLUS FACILE DE GERER UN GRAND NOMBRE DE REQUETES SIMULTANES GRACE A L'ASYNCHRONE

## R

### POINTS FORTS



LANGAGE PUISSANT POUR FAIRE DE L'ANALYSE DE DATAS



GRANDE VARIÉTÉ D'OUTILS POUR EFFECTUER DES OPÉRATIONS COMPLEXES SUR LA DATA

## LIBRAIRIES LES PLUS POPULAIRES SUR GITHUB

Nombre total de repositories matchant le nom de chaque librairie

