

Approfondimento: WEB SCRAPING

Pensa a quanto può essere utile raccogliere grandi quantità di dati da pagine web, per poi analizzarli o utilizzarli nei tuoi progetti, senza doverlo fare manualmente, ma con l'ausilio di semplici programmi che forniscono istruzioni informatiche. Questo è il **web scraping**, che significa letteralmente “**raschiare**” informazioni dal web in modo automatico.

Cos'è e come funziona il web scraping

Sono tanti i campi di applicazione del web scraping: dalla **raccolta di dati da varie fonti online** come articoli, prezzi di prodotti, recensioni, all'**acquisizione di informazioni aggiornate** su cui fare analisi approfondite, al **monitoraggio** dell'andamento dei prezzi di mercato o altri dati a fini statistici.

Per capire meglio come funziona lo scraping, pensa ai tre passi principali necessari per attuarlo: 1) visitare il sito web che ti interessa analizzare e attivare uno “**scraper**”, cioè il programma dove andrai a inserire le istruzioni di scraping; 2) raccogliere le informazioni che lo scraper trova in modo automatico grazie alle indicazioni date in precedenza; 3) copiare e salvare in un file o un database i dati raccolti.

Ci sono diverse applicazioni che puoi usare per iniziare in modo semplice a fare scraping, come [Google Sheets](#), che permette di estrarre dati da pagine web usando la funzione *ImportXML*, [Octoparse](#), un tool che consente di fare scraping anche senza saper programmare, oppure [ParseHub](#), un'altra opzione facile da usare per progetti semplici.

Se vuoi invece cimentarti nella scrittura di codici con linguaggio Python per fare scraping di pagine web, puoi usare [BeautifulSoup](#) oppure [Scrapy](#), quest'ultimo pensato per gestire grandi quantità di dati.

Alcuni esempi pratici di scraping

Vediamo insieme alcuni esempi di come il web scraping può essere applicato nella vita di tutti i giorni. Supponiamo che voglia comprare un paio di scarpe o un nuovo telefono, ma vuoi aspettare che il prezzo scenda. Puoi creare uno scraper che controlla ogni giorno i prezzi su vari siti web di shopping e ti avvisa quando il prezzo scende sotto una certa cifra:

```
python

if prezzo_corrente < prezzo_desiderato:
    invia_notifica("Il prezzo è sceso! È ora di comprare.")
```

Puoi usare lo scraping anche per raccogliere le previsioni meteo da diversi siti e creare un riepilogo delle previsioni per la tua zona. In questo modo, avrai sempre un'idea chiara di che tempo farà nei prossimi giorni...

```
python
```

```
for sito_meteo in siti_meteo:  
    meteo = scraper_meteo(sito_meteo)  
    aggiungi_alla_lista(meteo)
```

Se ti piace cucinare, puoi creare uno scraper che raccoglie ricette su un ingrediente specifico da diversi siti di cucina. Per esempio, se hai delle mele e non sai cosa cucinare, lo scraper può trovare tutte le ricette che includono mele e organizzarle per te:

```
python
```

```
for sito_ricette in siti_ricette:  
    ricette = scraper_ricette(sito_ricette, "mele")  
    aggiungi_alla_lista(ricette)
```

Se sei alla ricerca di un lavoro part-time o di uno stage, puoi creare uno scraper che visita diversi siti di offerte di lavoro e raccoglie tutte le opportunità che corrispondono ai tuoi interessi. Puoi anche programmare il tutto per inviarti una notifica quando ci sono nuove offerte:

```
python
```

```
if nuove_offerte_lavoro():  
    invia_notifica("Ci sono nuove offerte di lavoro disponibili!")
```

Ci sono molti **tutorial online** che ti possono insegnare a fare i primi esperimenti di scraping: fai una semplice ricerca con le parole "tutorial" e "scraping"!

Fare scraping per ASOC e regole di utilizzo

Il web scraping può rivelarsi **uno strumento molto potente anche per raccogliere dati e informazioni sul progetto che stai monitorando nel tuo percorso ASOC**. Con un po' di esercizio, potrai essere in grado di scrivere codici di scraping per estrarre informazioni da siti web istituzionali o di organizzazioni non profit collegate al progetto che stai osservando, e usare i dati raccolti per confrontare i risultati con le esigenze della tua comunità locale.

In tutti i casi, **è sempre importante rispettare le regole dei siti web**. Non tutti i siti, infatti, permettono lo scraping, perciò è importante leggere i termini di servizio del sito prima di iniziare. In generale, il web scraping è consentito quando i dati estratti sono dati pubblicamente disponibili e le informazioni raccolte non sono protette da un login. Fare un **uso responsabile del web scraping** richiede prudenza circa i termini di servizio del sito che si sta analizzando, i dati protetti da copyright e i dati personali, che di solito sono tutelati dalle leggi sulla privacy.

In conclusione...

Che si tratti di raccogliere dati per la tua ricerca di monitoraggio civico, trovare le migliori offerte o semplicemente rendere la tua vita più organizzata, lo scraping ti permette di **ottenere le informazioni che ti servono in modo semplice e veloce**. Non importa se sei alle prime armi, ci sono strumenti e risorse per tutti. Quindi, rimboccati le maniche, scegli uno degli strumenti di cui abbiamo parlato, e inizia a “raschiare” – sempre in modo responsabile! – le informazioni dal web.