

# PROYECTO TECNOLÓGICO INTEGRADOR

---

Integrantes:

Duarte, Carolina Alejandra

Ferreya, Santiago Nicolás

Martins, Cesar Lucas

Mathieu, Ezequiel

Trejo, Ernesto Gaston

Urcola, Ma. Victoria

---

Cohorte 2023



# Scraping

---

El web scraping, o raspado web, es una técnica que permite adquirir información estructurada desde un sitio web. Esto se logra mediante la simulación de la navegación humana, con el propósito de realizar análisis de datos a gran escala (Big Data) o automatizar procesos web.

A nivel mundial, diversas industrias y profesionales hacen uso del scraping, incluyendo:

- **Empresas:** Lo utilizan para recopilar datos competitivos, analizar el mercado, rastrear precios de productos y obtener información sobre clientes y usuarios.
- **Investigadores:** Esta técnica es fundamental para obtener datos relevantes en investigaciones académicas y científicas en diversas disciplinas.
- **Periodistas:** Los periodistas la emplean para acceder a información pública en línea y realizar investigaciones periodísticas.
- **Desarrolladores:** Usan el scraping para crear aplicaciones web que agregan información de múltiples fuentes.
- **Gobierno:** Algunos organismos gubernamentales recurren al scraping para recopilar datos públicos y mejorar la toma de decisiones.

El web scraping se utiliza ampliamente para obtener grandes volúmenes de información de forma automatizada. Sus principales aplicaciones y ventajas incluyen:

- **Análisis Competitivo:** Las empresas pueden analizar a sus competidores, ajustar estrategias de marketing, mejorar el SEO y tomar decisiones comerciales fundamentadas.
-



# Scraping

---

- Automatización de Tareas: Permite la automatización de tareas como reservas en línea, agregación de datos bancarios, procesos de pago y verificación de servicios web.
- Optimización de Marketing y SEO: Facilita la comprensión del mercado y mejora los recursos dedicados al marketing y al SEO.
- Toma de Decisiones Basada en Datos: Proporciona datos actualizados y relevantes para tomar decisiones informadas.

La legalidad del scraping web depende de factores como la disponibilidad pública de los datos y la protección por inicio de sesión. En Argentina, la Ley de Protección de Datos Personales (Ley 25.326) regula el scraping de datos personales, requiriendo consentimiento adecuado. También se deben respetar leyes de derechos de autor y términos y condiciones de los sitios web.

## Lenguajes de Programación y Librerías Relevantes:

- Python: Es altamente recomendado debido a su facilidad de uso. Librerías populares incluyen BeautifulSoup para extracción de datos personalizable y Scrapy para descargas masivas.
  - JavaScript: Útil para sitios web con interfaces dinámicas, aunque menos adecuado para extracciones repetitivas.
  - Node.js: Ideal para sitios avanzados con contenido dinámico.
  - C++: Aunque eficiente, se evita en negocios debido a la complejidad.
  - PHP: Efectivo para extraer datos y programar tareas en múltiples sitios web.
-



# Scraping

---

## Ejemplos Prácticos de Aplicación

- Comparación de Precios en Comercio Electrónico: Monitorizar y comparar precios de productos en diferentes tiendas en línea.
- Análisis de Noticias Financieras: Recopilar noticias y comentarios sobre eventos económicos para la toma de decisiones financieras.
- Gestión de Reputación en Redes Sociales: Recopilar comentarios y reseñas de usuarios en redes sociales para evaluar la percepción de la marca en línea.

## Ventajas de Web Crawling y Data Scraping:

- Reducción de la carga de trabajo y costos de personal.
- Aumento significativo de la velocidad de los procesos.
- Minimización del error humano en la recopilación de datos.
- Capacidad para manejar grandes volúmenes de información.
- Obtención de datos en formatos procesables y listos para su análisis.

## Bibliografía

<https://www.amnislabs.com/web-scraping>.  
<https://datstrats.com/blog/que-es-scraping-usos-aplicaciones/>  
<https://www.iubenda.com/es/help/111939-es-legal-el-web-scraping-lo-que-hay-que-saber>  
<https://www.areacucuta.com/que-lenguajes-son-empleados-en-web-scraping/>

---



