EN LA PESCA DE DATOS, NECESITAS UN PESCADOR PROFESIONAL

Encontrar el valor en un océano de información

Aralí Valderrama, Jorge Marcial Garcia Rizo

La información es poder. Tener la habilidad de obtener la información correcta para la toma de decisiones es crucial para cualquier empresa. Afortunadamente, existen tecnologías avanzadas de Web scraping que te pueden ayudar a obtener esta información. La clave está en tener un guía experto para asegurar que tu "pesca" de datos sea fructífera.

Desde el nacimiento del internet, en 1991, los humanos hemos creado cantidades inimaginables de datos. ¿Quiénes somos? ¿Qué edad tenemos? ¿Qué nos gusta? ¿Dónde vivimos? ¿Qué lugares frecuentamos? ¿Qué opinión tenemos sobre cierto tema? ¿Cuánto gastamos en entretenimiento? ¿Dónde nos gusta aprender cosas nuevas? Todos estos datos, están disponibles en el internet, de manera agregada por millones de personas.

Según **cálculos del 2019, existen más de 5.85 billones de sitios de internet. Si hablamos en términos de tráfico, o la cantidad de datos que se mueven en el internet, en 2019 llegamos a 2 zettabytes por año. Un zettabyte es el equivalente a 36,000 años de video en alta definición, o a transmitir todo el catálogo de Netflix 3,177 veces. En fin, el internet es inimaginablemente grande, como el océano.**

Y como en el océano, podemos aprovechar sus productos para hacer negocios. Con la información que tenemos en internet, podemos aprender muchísimo sobre nuestros clientes, sus preferencias, opiniones, gustos, lugares favoritos, etc. Todas las empresas ofrecen un producto o servicio al mercado, y para ello es imperativo conocer a fondo a sus clientes y competidores. Y toda esta información está ahí, en ese inmenso océano llamado Internet.

Pero ¿cómo podemos "pescar" esos datos? El océano es tan grande que no sabríamos ni por dónde empezar. Consigues una lancha, una gran red y te lanzas al mar… tiras tu red, esperas un tiempo y recoges tus peces. De seguro vas a obtener peces chicos, grandes, e incluso bastante basura. ¿Qué tal si tu quisieras pescar dorado específicamente? De seguro tendrías que ir con un pescador con experiencia, que te podría decir que estos peces viven en el Pacífico y que generalmente se encuentran a una profundidad de 10 metros. El pescador te indicaría qué tipo de malla es la más conveniente para esta pesca, y te podría decir en qué horarios tendrás más probabilidad de encontrarlos, en una ubicación específica. En esta pesca, de seguro conseguirás muchos dorados, y mucho menos basura.

Copyright Derechos Reservados Sistemas Interactivos de Consultoría, S.A. de C.V.

El ***Web scraping***, llevado por un experto, funciona más o menos de la misma forma. **Un pescador de datos con experiencia te puede ayudar a definir el tipo de datos que necesitas "pescar" para resolver un problema específico en tu negocio.** Después puede programar una herramienta para que vaya y busque esos datos específicos en el océano del internet. Finalmente, te puede ayudar a analizar estos datos para extraigas la información que necesitas para tomar decisiones en tu negocio.

El ***Web scraping*** tiene un sinfín de aplicaciones, en todas las industrias. Con los datos que se "pescan" puedes saber cuáles son los artículos más leídos de ciertas publicaciones, revisar la reputación de tu empresa y sus competidores en las redes sociales, cazar tendencias (coolhunting) para hacer campañas de marketing dirigidas específicamente a los gustos de tu mercado, monitorear los precios de tus competidores para desarrollar estrategias de pricing más agresivas, etc.

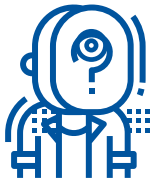
Hacer Web scraping es más sencillo de lo que piensas, y como todas las nuevas tecnologías, está más cerca de tu alcance que nunca. Principalmente se necesitan tres elementos:

Lenguaje de programación

Ésta es la forma en que le dices a las computadoras qué información debe buscar y dónde. El código que se genera captura de manera inteligente los datos que quieres encontrar, y las reglas del juego para hacerlo. Esta programación se puede hacer o cambiar cuantas veces quieras, dependiendo de la información que quieras buscar en un determinado momento.

Infraestructura para escalar

Una sola computadora, por más poderosa que sea, no podrá abarcar tantos datos como son necesarios en algunos casos. Para ello se necesitan varios procesadores con un desarrollo específico que pueda manejar big data y analizarla con machine learning.



Anonimato

Por todas las reglas de privacidad vigentes en los diferentes países, no se debe usar datos específicos de un usuario o un sitio en particular. Sin embargo, si se hace de manera agregada (millones de datos de usuarios anónimos) esto no es un problema. Lo importante es estar bien asesorado sobre el cuidado de los datos que se van a obtener.

Hace algunos años el ***Web scraping*** era una herramienta novedosa que pocos utilizaban. Hoy es una necesidad para la mayoría de los negocios porque sus competidores de seguro ya lo están utilizando. Probablemente tú ya has tenido contacto con ***Web scraping***. Si alguna vez has buscado precios de boletos de avión o de noches de hotel de forma comparativa existen sitios como kayak.com o trivago.com. Estos sitios hacen ***Web scraping*** para que el consumidor pueda tomar la decisión más conveniente, con toda la información disponible.

Otros usos comunes son, en la industria de retail, monitorear los precios de la competencia por SKU para determinar el precio de productos iguales o similares en tu negocio. En el sector financiero, se pueden obtener datos para comparar el margen de los productos de seguros. Incluso en el mundo de la política se puede utilizar para monitorear la reputación de un candidato o persona pública.

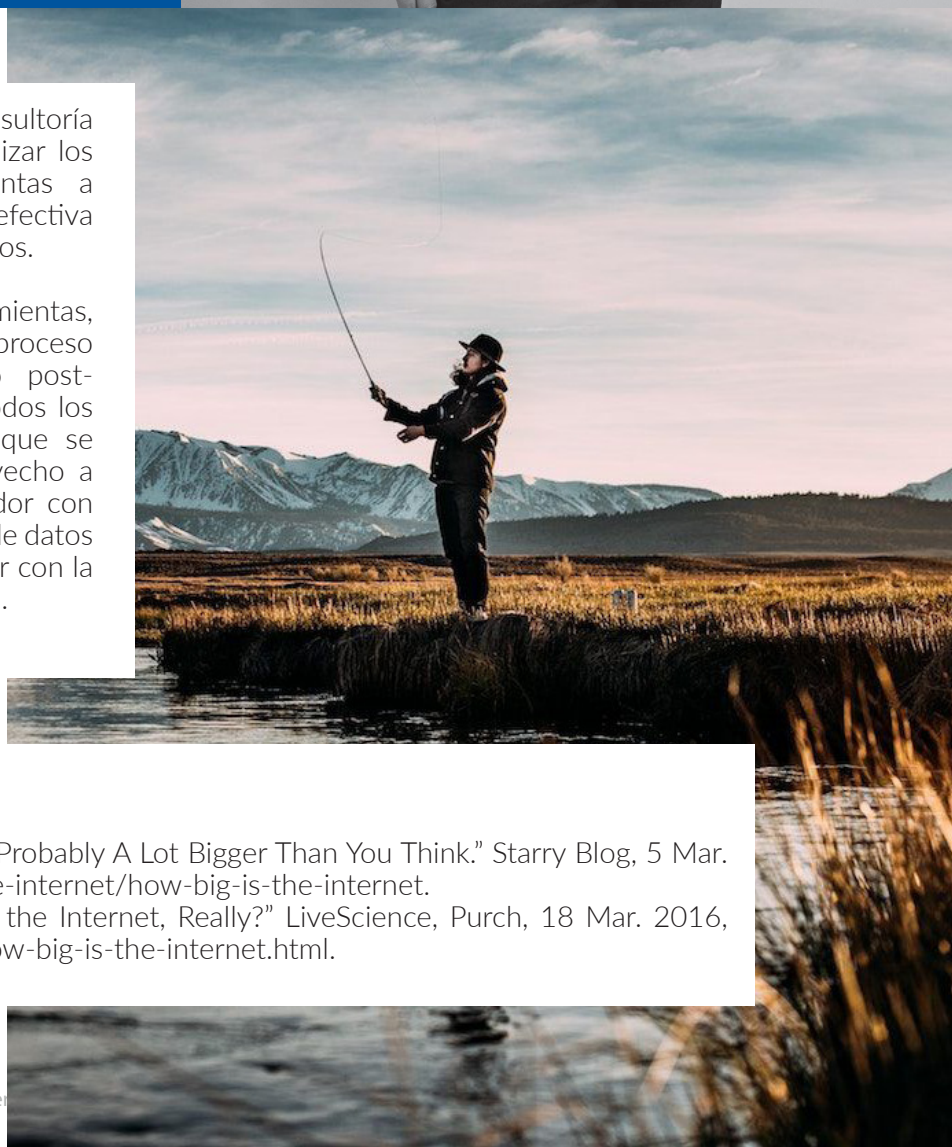
Copyright Derechos Reservados Sistemas Interactivos de Consultoría, S.A. de C.V.

En esta era en que la información es poder, no puedes quedarte atrás en la "pesca" de datos para tomar decisiones. Sintec puede ser ese pescador con experiencia que puede guiarte si quieres embarcarte en la pesca de *Web scraping*.

ARALÍ
VALDERRAMA,
Directora Sintec Digital

Gracias a nuestra experiencia en consultoría de más de 30 años, podemos analizar los problemas, determinar las preguntas a resolver, y dirigir una estrategia efectiva para cumplir tus objetivos estratégicos.

Sintec no solo implementa herramientas, sino que te acompaña en todo el proceso desde diagnóstico hasta servicio post-implementación, llevando a cabo todos los cambios en procesos y prácticas que se necesitan para sacar el mejor provecho a estas herramientas. Con un pescador con experiencia como Sintec, el océano de datos ya no es tan grande, y es fácil contar con la información que tu negocio necesita.



Referencias:

- "How Big Is The Internet? Hint: Probably A Lot Bigger Than You Think." Starry Blog, 5 Mar. 2020 starry.com/blog/inside-the-internet/how-big-is-the-internet.
- Pappas, Stephanie. "How Big Is the Internet, Really?" LiveScience, Purch, 18 Mar. 2016, www.livescience.com/54094-how-big-is-the-internet.html.

Copyright Derechos Reser

Acerca de los autores

Aralí Valderrama
Directora de Sintec Digital, CDMX
arali.valderrama@sintec.com

Con más de 25 años de experiencia profesional, ha desarrollado habilidades clave para la innovación y la excelencia en minería de datos, planeación estratégica, logística y calidad. Como Directora de Sintec Digital, está orientada a la aplicación lógica, resolución de problemas de forma analítica y con un enfoque disruptivo.

Jorge Marcial García Rizo
Consultor Sintec Digital, Monterrey
jorge.garcia@sintec.com

Con más de 6 años de experiencia profesional, ha participado en el desarrollo de software de producto y software como servicio, para diversas empresas y startups. Como parte del equipo de Sintec Digital, está orientado al desarrollo de soluciones de software que impulsan las soluciones diseñadas por las prácticas de business consulting con características de alta escalabilidad, replicación y disponibilidad.

CIUDAD DE MÉXICO / MONTERREY / BOGOTÁ / SÃO PAULO / SANTIAGO DE CHILE

www.sintec.com
informes@sintec.com

/SintecConsulting @Sintec @Sintec_