

Herramientas de recolección de datos en la era digital y sus retos en materia de privacidad y protección de datos personales

Por: Santiago Rafael Torres Zapata & Catalina Chaparro Casas

Dado el apogeo de la digitalización de las relaciones comerciales y la necesidad de acceder a información masiva de los usuarios de una manera más rápida, con el objetivo de ofrecer productos y prestar servicios de forma eficiente, existe una tendencia a utilizar mecanismos de recolección que implican retos de cara a la privacidad y la protección de los datos personales, entendidos en palabras de Martín Hilbert como “el *“nuevo petróleo” de la industria digital*”.

La privacidad y la protección de los datos personales se cataloga como un derecho inalienable al ser humano como titular de los datos y, por ende, impone un sin número de obligaciones a quienes traten dichos datos, ya sea que éstos sean responsables o meros encargados, las cuales se pueden encontrar en la Ley 1581 de 2012, el Decreto 1377 de 2013 y demás normas aplicables.

El tratamiento de datos en la era digital se ha venido adecuando a los nuevos modelos de negocio, de manera que cada vez existen más medios que permiten llevar a cabo una recolección de información a través del Internet y, con éstos, se han incrementado las exigencias en el cumplimiento de las obligaciones a cargo de la persona, sea esta natural o jurídica, que le esté dando un tratamiento a los datos de los usuarios de Internet.

Así las cosas, adentrándonos en el tratamiento de datos, en un mundo digitalizado, podemos encontrar algunas de las herramientas para el almacenamiento, recuperación y/o recolección de información en equipos de los usuarios de Internet como lo son el uso de las *cookies*, el *web scraping* y el *web crawling* que, en general, son utilizadas para identificar personas, elaborar sus perfiles y adaptar el funcionamiento del sitio web conforme a dichos perfiles.

Las *cookies* son un paquete de datos y se refieren a los archivos que se almacenan en cualquier equipo terminal, sea este un computador, un celular, una Tablet, entre otros. Éstas son creadas por cada sitio web al que el usuario accede a través de Internet y pretenden realizar un seguimiento de la actividad de un usuario en la Web pues permiten que los diferentes prestadores de servicios en Internet conozcan en detalle los hábitos de navegación del usuario y puedan utilizarlos como ventaja para su oferta de servicios, como base de publicidad y/o como oportunidad de mejora en productos y servicios.

Sin embargo, la utilización de cookies requiere el cumplimiento de obligaciones a cargo de los prestadores de servicios de Internet, de manera que, es importante la asesoría tanto técnica como jurídica con el fin de llevar a cabo lo siguiente:

- (i) Implementar un sistema de fácil acceso mediante el cual el usuario pueda conocer de manera clara y completa la información sobre la utilización de las *cookies*, el tratamiento y las finalidades que se le darán a sus datos almacenados.
- (ii) Esto incluye la obtención de la debida autorización por parte de los usuarios de Internet que permita el tratamiento de sus datos para los fines incluidos en la política de *cookies*.

- (iii) Poner en conocimiento de los usuarios los derechos y opciones disponibles para la gestión de sus *cookies*.

De esta forma se garantiza la transparencia exigida por la Ley para el uso de las *cookies* con el fin de llevar a cabo la recolección de información de usuarios en Internet.

Por otro lado, el *web scraping* corresponde a la utilización de un software para la extracción y recolección de información de diferentes páginas web, de manera automática y masiva, incluyendo los datos de sus usuarios. Esta técnica permite alimentar bases de datos, hacer una migración de un sitio web, monitorear precios de la competencia, generar alertas de cambios de páginas, fichas de productos, entre otras finalidades, por lo que podría ser considerada como la fuente más importante de los datos.

Si bien es una buena herramienta permitida para recopilar información masiva, no siempre es posible acceder a todos los sitios web. Esto, por cuanto algunos sitios, con el fin de proteger la información y datos de los cuales son responsables, deciden restringir o bloquear su acceso, mediante mecanismos que identifican este tipo de accesos automatizados cuando no provienen de un ser humano. Otra desventaja del *web scraping* es que, al ser un mecanismo automatizado, un simple cambio en la página web por parte de su titular y de la cual se quiere recolectar la información, puede generar errores en dicho proceso de recolección.

Para finalizar, muy similar al *web scraping*, se encuentra la técnica de *web crawling* la cual utiliza programas que navegan y analizan los sitios web en Internet, con el fin de recolectar información que ha sido ordenada y programada previamente, clasificarla, indexarla y poder crear una propia base de datos a partir de ella. Mediante esta herramienta, se pueden, por ejemplo, comparar precios de productos en los diferentes portales, recolectar correos electrónicos y direcciones de las empresas e indexar la información recolectada en otros buscadores.

A pesar de que es un mecanismo utilizado para el beneficio de las empresas prestadoras de servicio en Internet, con el fin de aparecer en una mayor cantidad de búsquedas, esta técnica genera el riesgo de un inadecuado uso de datos de los usuarios por parte de terceros, motivo por el cual, algunos prestadores de servicios también prefieren bloquear este tipo de actividades, mas por cuanto exponen la información de sus usuarios y ello puede comprometer la responsabilidad de dichas empresas.

Por lo expuesto, dado los riesgos que presenta la utilización de cualquiera de estas técnicas en materia de protección de datos y privacidad de los usuarios por parte de prestadores de servicios en Internet, es importante: (i) identificar las implicaciones de la utilización de determinado mecanismo tecnológico para la recolección de información, (ii) establecer los procesos, medidas y documentos que se deben adoptar para dar cumplimiento a la normativa sobre esta materia e (iii) implementar estrategias de protección de la información de sus usuarios, por ejemplo contra el *web scraping* y *web crawling*. Lo anterior, con el propósito de salvaguardar la privacidad y la protección de los datos, garantizando la confianza de los usuarios en la red y la transparencia en el desarrollo de cualquier actividad comercial que requiera una recolección de información.