

Cloudflare hace historia bloqueando las consultas de las IA y sienta las bases para transformar el negocio

Por: Kate Knibbs. 06/07/2025

La extracción de información de páginas web con IA puede estar llegando a su fin. Al menos si Cloudflare se sale con la suya.

El año pasado, la empresa de infraestructuras de internet Cloudflare lanzó herramientas que permitían a sus clientes bloquear a los *scrapers* (raspadores web) de IA. Hoy, la empresa ha dado un paso más en su lucha contra el *scraping* sin permisos. **Ha pasado a bloquear los rastreadores de IA por defecto para sus clientes y está avanzando con un programa de pago por rastreo que permite a los clientes cobrar a las empresas de IA por rastrear sus sitios web.**

Los raspadores de datos llevan décadas buscando información en internet

Sin ellos, la gente perdería herramientas en línea de vital importancia, desde [Google Search](#) hasta la inestimable labor de preservación digital de [Internet Archive](#). Pero el auge de la IA ha producido un auge correspondiente de raspadores web centrados en la IA, y estos robots rastrean páginas web con una frecuencia que puede imitar un ataque DDoS, sobrecargando los servidores y dejando fuera de servicio los sitios web. Incluso cuando los sitios web pueden soportar el aumento de actividad, muchos no quieren que los *scrapers* rastreen su contenido, especialmente las publicaciones de noticias que exigen a las empresas de inteligencia artificial que paguen por utilizar su trabajo. “Hemos estado tratando febrilmente de protegernos”, afirma Danielle Coffey, presidenta y CEO del grupo comercial News Media Alliance, que representa a varios miles de medios estadounidenses.

Hasta ahora, según explica a WIRED Will Allen, responsable de control de IA, privacidad y productos de medios de Cloudflare, más de un millón de sitios web de clientes han activado sus antiguas herramientas de bloqueo de [robots de IA](#). Ahora, millones más tendrán la opción de mantener el bloqueo de bots por defecto. Cloudflare también afirma que puede identificar incluso a los *scrapers* “en la sombra”

que no son publicitados por las empresas de IA. La empresa señaló que utiliza una combinación patentada de análisis de comportamiento, huellas dactilares y aprendizaje automático para clasificar y separar los bots de IA de los bots “buenos”.

El Protocolo de Exclusión de Robots (Robots Exclusion Protocol), una norma muy extendida en internet que se aplica a menudo a través de un archivo robots.txt, ayuda a los editores a bloquear los robots caso por caso, pero su cumplimiento no es obligatorio por ley, y hay muchas pruebas de que algunas empresas de IA intentan eludir los intentos de bloquear sus robots. “Robots.txt es ignorado”, afirma Coffey. Según un informe de la plataforma de licencias de contenidos Tollbit, que ofrece su propio mercado para que los editores negocien con las empresas de IA el acceso a los bots, el [“raspado de IA”](#) sigue en aumento, incluido aquel que ignora robots.txt. Tollbit descubrió que solo en marzo de 2025 más de 26 millones de *scrapes* ignoraron el protocolo.

El cambio de Cloudflare representa un obstáculo para los *scrapers*

El plan de la compañía podría dar a los editores más influencia para negociar, ya sea a través del programa Pay Per Crawl o de otra manera. “Esto podría cambiar radicalmente la dinámica de poder. Hasta ahora, las empresas de IA no tenían que pagar por las licencias de los contenidos, porque sabían que podían apropiarse de ellos sin consecuencias”, afirma Nicholas Thompson, CEO de *Atlantic*, y antiguo redactor jefe de WIRED. “Ahora tendrán que negociar, y se convertirá en una ventaja competitiva para las empresas de IA que puedan llegar a más y mejores acuerdos con más y mejores editores”.

La *startup* de IA ProRata, que opera el motor de búsqueda de IA Gist.AI, ha aceptado participar en el programa Pay Per Crawl, según su CEO y fundador, Bill Gross: “Creemos firmemente que todos los creadores de contenido y editores deben ser compensados cuando su contenido se utiliza en respuestas de IA”.

Por supuesto, queda por ver si las grandes empresas del sector de la inteligencia artificial participarán en un programa como Pay Per Crawl, que se encuentra en fase beta. Empresas como OpenAI han llegado a acuerdos de licencia con diversos socios editoriales, entre ellos [Condé Nast](#), la empresa matriz de WIRED, pero no se han revelado los detalles concretos de estos acuerdos, ni siquiera si el acuerdo

incluye el acceso a los bots.

Mientras tanto, existe todo un ecosistema en línea de tutoriales sobre cómo eludir las herramientas de bloqueo de bots de Cloudflare dirigidas a los raspadores web. A medida que se despliegue el bloqueo por defecto, es probable que estos esfuerzos continúen. **Cloudflare hace hincapié en que los clientes que quieran dejar que los robots hagan *scraping* sin impedimentos podrán desactivar la configuración de bloqueo.** “Todo el bloqueo es totalmente opcional y queda a discreción de cada usuario”, concluye Allen.

[LEER EL ARTÍCULO ORIGINAL PULSANDO AQUÍ](#)

Fotografía: Wired

Fecha de creación

2025/07/06