

Ayuda de Search Console

[SEARCH CONSOLE](#)[FORO](#)

Cómo bloquear URLs con robots.txt

Información sobre los archivos robots.txt

SIGUIENTE: [CREAR UN ARCHIVO ROBOTS.TXT](#)

Un archivo `robots.txt` es un archivo que se encuentra en la raíz de un sitio e indica a qué partes no quieres que accedan los rastreadores de los motores de búsqueda. El archivo utiliza el [Estándar de exclusión de robots](#), que es un protocolo con un pequeño conjunto de comandos que se puede utilizar para indicar el acceso al sitio web por sección y por tipos específicos de rastreadores web (como los rastreadores móviles o los rastreadores de ordenador).

¿Para qué sirve el archivo robots.txt?

Archivos que no son de imagen

Para los archivos que no son de imagen (es decir, las páginas web), el archivo robots.txt solo debe usarse para controlar el tráfico de rastreo. Normalmente, esto sirve para evitar que el rastreador de Google sobrecargue el servidor o para no gastar todo el presupuesto rastreando páginas poco importantes del sitio web o páginas que sean muy parecidas. **No uses robots.txt para ocultar las páginas web de los resultados de la Búsqueda de Google**, porque es posible que otras páginas dirijan a tu página; de esta forma, tu página podría indexarse evitando el archivo robots.txt. Si quieres bloquear tu página para que no aparezca en los resultados de búsqueda, usa un método distinto, como la protección mediante contraseña o las [directivas o etiquetas noindex](#).

Archivos de imagen

El archivo robots.txt **evita** que los archivos de imagen aparezcan en los resultados de la Búsqueda de Google. (Sin embargo, no evita que otras páginas u otros usuarios enlacen con tu imagen.)

Archivos de recursos

Puedes usar el archivo robots.txt para bloquear los archivos de recursos (como los de imagen, de script o de estilo que no sean importantes) si crees que la pérdida no afectará de forma importante a las páginas que se carguen sin dichos recursos. No obstante, si la ausencia de los recursos complica el análisis del rastreador de Google, no deberías bloquearlos; de lo contrario, Google no podrá analizar correctamente las páginas que dependan de dichos recursos.

Limitaciones del archivo robots.txt

Antes de crear un archivo `robots.txt`, deberías saber qué riesgos conlleva este método de bloqueo de las URL. Es posible que en algún momento quieras utilizar otros mecanismos para garantizar que tus URL no se pueden encontrar en la Web.

- **Las instrucciones de robots.txt son solo indicaciones**

Las instrucciones de los archivos `robots.txt` no pueden forzar el comportamiento de los rastreadores en el sitio, sino que sirven de indicaciones para los rastreadores que acceden a él. Mientras que Googlebot y otros rastreadores web respetables obedecen las instrucciones de un archivo `robots.txt`, puede que otros rastreadores no lo hagan. Por lo tanto, si quieres proteger la información de los rastreadores web, es mejor que utilices otros métodos de bloqueo, como [proteger los archivos privados con contraseña en el servidor](#).

- **Los distintos rastreadores interpretan la sintaxis de varias formas**

Aunque los rastreadores web respetables siguen las indicaciones de los archivos `robots.txt`, cada rastreador puede interpretar las indicaciones de forma distinta. Debes conocer la sintaxis adecuada para dirigirte a los diferentes rastreadores web, ya que puede que algunos no entiendan determinadas instrucciones.

- **Las indicaciones de robots.txt no pueden impedir las referencias a las URL que provienen de otros sitios**

Mientras que Google no rastreará ni indexará el contenido bloqueado por `robots.txt`, es posible que encontremos e indexemos una URL no autorizada de otras partes de la Web. Por lo tanto, es posible que la dirección URL y otra información de disponibilidad pública, como el texto del enlace que lleva al sitio web, todavía puedan aparecer en los resultados de la Búsqueda de Google. Puedes evitar que tu URL aparezca en los resultados de la Búsqueda de Google completamente mediante otros métodos de bloqueo de URL, como la [protección de los archivos mediante contraseña en el servidor](#) o el [uso de la cabecera de respuesta o la metaetiqueta noindex](#).

Nota: La combinación de varias indicaciones de rastreo e indexación puede hacer que algunas indicaciones bloqueen otras. Para obtener más información sobre cómo configurar estas indicaciones correctamente, lee la sección [Combinar el rastreo con las indicaciones de indexación o publicación](#) de la documentación de Google Developers.

SIGUIENTE: CREAR UN ARCHIVO ROBOTS.TXT

¿Te ha sido útil este artículo?

SÍ

NO

Cómo bloquear URLs con robots.txt

Información sobre los archivos robots.txt

Crear un archivo robots.txt

Probar tu archivo robots.txt con el Probador de robots.txt

Enviar el archivo robots.txt actualizado a Google

esp... 

©2016 Google - Política de privacidad - Condiciones del servicio 