

Eliminación de contenidos indexados en buscadores

27/07/2024 - 02:42:46

[Imprimir el artículo FAQ](#)

Categoría:	Alojamiento web	Última Actualización:	Mié, 16 May 2018 - 09:35:50
Estado:	public (all)		

Palabras Clave

eliminar contenido indexado indexados contenidos buscadores motores de búsqueda Google

Solución (público)

En ocasiones un buscador como Google puede estar mostrando contenidos en su resultado y caché que nos gustaría que fuese eliminado. En primer lugar hay que entender cómo funcionan los buscadores.

Los motores de búsqueda almacenan información de un gran número de páginas que extraen de Internet. Para obtener la información de dichas páginas usan unos programas autónomos denominados arañas o crawlers, que son los encargados tanto de obtener los contenidos - como si de una conexión de usuario se tratase - así como de ir siguiendo los enlaces que pueda encontrar en dicha página.

Si por un error, un documento confidencial estuviese disponible durante unos días en un sitio web, podría ser indexado por uno de estos sistemas y comenzaría a aparecer tanto en los resultados del buscador como en sus contenidos de cache, que permite acceder a los contenidos de un sitio web aunque este no esté disponible.

Para lograr que una web o un contenido desaparezca de los resultados y la cache de un buscador, tendremos que seguir este procedimiento.

Primero

Debemos eliminar el contenido original en nuestro servidor, o bloquear el acceso al mismo para evitar que en el futuro vuelva a ser indexado. El bloqueo en el acceso a un contenido (ej. fichero html o pdf) se puede realizar de diferentes formas, para conocer más sobre cómo hacerlo se puede consultar la [1]siguiente dirección o en el caso de que se trate de un contenido que no sea un fichero HTML como un archivo PDF o un archivo de imagen puede acudir a la [2]siguiente dirección, en la sección "Implementación práctica de X-Robots-Tag con Apache" de dicha página. Sin otra intervención un contenido que sea eliminado o bloqueado en el acceso terminará desapareciendo de los resultados y caché del buscador pasado un tiempo en el que el crawler vea que ese contenido ya no existe.

Segundo

Si deseamos acelerar el proceso de eliminación en resultados y en cache, tendremos que seguir los siguientes pasos.

Hemos de demostrar frente a Google la propiedad sobre el dominio en el que se aloja el contenido a eliminar. Para esto nos dirigiremos al sitio de Google Search Console y accederemos al mismo con una cuenta de Gmail o @unizar.es que tenga activado el acceso a Google Apps.

Indicaremos que deseamos Añadir una propiedad y a continuación nos solicitará el nombre del dominio del que somos dueños. Ejemplo: tecnomap.unizar.es

A continuación nos pedirá que usemos uno de los difentes métodos disponibles para demostrar esa propiedad, que incluyen subir un fichero especial al alojamiento, modificar la página principal del alojamiento con una etiqueta <meta> específica, y otros métodos alternativos.

Una vez cumplido el requisito de demostrar la propiedad de dicho dominio, pulsando en Verificar podremos acceder al panel de administración Search Console para nuestro dominio.

Este panel, además de otras muchas funcionalidades, nos permite solicitar a Google la eliminación de contenidos de sus resultados y caché. Para esto nos dirigiremos al menú Índice de Google-> Eliminación de URL -> Ocultar temporalmente, y en ese momento indicaremos la URL completa que deseamos que Google deje de mostrar en sus resultados y caché. Este proceso suele tardar unas 24 horas en ser efectivo, pero esta eliminación de resultados durará unos 90 días, pasado ese tiempo si el archivo con dicho contenido sigue siendo accesible a los crawlers volverá a ser indexado y mostrado salvo que se trabaje en el punto primero de esta FAQ

Para más información en relación con este problema se puede visitar la web de ayuda de Google - [3]Eliminar Información de Google

[1] <https://support.google.com/webmasters/topic/4598466>

[2] https://developers.google.com/webmasters/control-crawl-index/docs/robots_meta_tag?hl=es-419

[3] <https://support.google.com/webmasters/answer/6332384?hl=es>