# SITEMAP

Un sitemap es un tipo de tabla de contenido para tu web. Un mapa del sitio que enumera las URLs de las subpáginas de un dominio. El objetivo es que Google u otros motores de búsqueda representen la estructura de la página.

Un mapa del sitio también puede contener metadatos importantes que indican a los rastreadores de los motores de búsqueda cuándo se actualizó una página por última vez, con qué frecuencia se modifica y cómo se relaciona cada página con otras páginas.

Es necesario cuando:

* El sitio web es muy extenso. En este caso, es más probable que los rastreadores web de Google pasen por alto algunas de las páginas nuevas o actualizadas recientemente mientras rastrean.
* El sitio web tiene un gran archivo de páginas de contenido que están separadas o no relacionadas. Los enlaces internos apenas existen. Se puede enumerar estas páginas en un mapa del sitio para que Google no las pase por alto.
* El sitio web es completamente nuevo y hasta el momento hay muy pocos enlaces externos que lo apuntan. Googlebot y otros rastreadores web rastrean la web siguiendo los enlaces de una página a otra. Google y compañía pueden no rastrear una página que carezca de vínculos de retroceso.
* El sitio utiliza contenido multimedia enriquecido, aparece en Google News o incluye otras anotaciones compatibles con Sitemaps. Google puede considerar información adicional de Sitemaps si es necesario.

### Mapa del sitio formatos

Hay varias formas de crear un mapa del sitio, así como varios formatos de mapa del sitio que los motores de búsqueda pueden utilizar para comenzar.

Vamos a aprender uno de ellos: sitemap.xml

#### Sitemap.xml

1. Para su creación automática: [Create your Google Sitemap Online - XML Sitemaps Generator (xml-sitemaps.com)](https://www.xml-sitemaps.com/)
2. Manual

Ejemplo de un Sitemap XML:



Para obtener más información sobre Sitemaps: <https://www.sitemaps.org/protocol.html>

### FORMATO XML DE SITEMAPS

El formato del protocolo Sitemap consta de etiquetas XML. El archivo debe estar codificado en UTF-8.

El Sitemap debe:

* Comenzar con una etiqueta de apertura <urlset> y terminar con una de cierre </urlset>.
* Especificar el espacio de nombres (protocolo estándar) en la etiqueta urlset.
* Incluir una entrada <[url](https://www.sitemaps.org/es/protocol.html#urldef)> para cada dirección URL como una etiqueta XML principal.
* Incluir una entrada secundaria <[loc](https://www.sitemaps.org/es/protocol.html#locdef)> para cada etiqueta principal <url>.

Las demás etiquetas son opcionales. La compatibilidad de estas etiquetas opcionales puede variar en función del motor de búsqueda.

Además, todas las direcciones URL de un Sitemap deben proceder de un único host, como www.example.com ó store.example.com.

#### Ejemplo de un Sitemap en formato XML

El siguiente ejemplo muestra un Sitemap que contiene únicamente una URL y usa todas las etiquetas opcionales. Las etiquetas opcionales están en cursiva.

<?xml version="1.0" encoding="UTF-8"?><

<[urlset](https://www.sitemaps.org/es/protocol.html#urlsetdef) xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">

<[url](https://www.sitemaps.org/es/protocol.html#urldef)>

<[loc](https://www.sitemaps.org/es/protocol.html#locdef)>http://www.example.com/</loc>

*<*[*lastmod*](https://www.sitemaps.org/es/protocol.html#lastmoddef)*>2005-01-01</lastmod>*

*<*[*changefreq*](https://www.sitemaps.org/es/protocol.html#changefreqdef)*>monthly</changefreq>*

*<*[*priority*](https://www.sitemaps.org/es/protocol.html#prioritydef)*>0.8</priority>*

    </url>

</urlset>

### Definiciones de las etiquetas XML

Las etiquetas XML disponibles se describen a continuación.

|  |  |  |
| --- | --- | --- |
| **Atributo** |  | **Descripción** |
| <urlset> | **obligatorio** | Encapsula el archivo y hace referencia al protocolo estándar actual. |
| <url> | **obligatorio** | Etiqueta principal de cada entrada de URL. Las demás etiquetas son secundarias de esa. |
| <loc> | **obligatorio** | URL de la página. Esta URL debe comenzar con el protocolo (por ej., http) y acabar con una barra diagonal, si su servidor web así lo requiere. Este valor debe contener menos de 2.048 caracteres. |
| <lastmod> | opcional | Fecha de la última modificación del archivo. Esta fecha debe encontrarse en formato [Fecha y hora de W3C](http://www.w3.org/TR/NOTE-datetime). Este formato le permite omitir la parte referente a la hora, si así lo desea, y utilizar AAAA-MM-DD.  Tenga en cuenta que esta etiqueta es independiente de la cabecera “If-Modified-Since (304)” que puede mostrar el servidor y que los motores de búsqueda pueden utilizar la información de ambas fuentes de forma diferente. |
| <changefreq> | opcional | Frecuencia con la que puede cambiar esta página. Este valor proporciona información general a los motores de búsqueda y es posible que no se corresponda exactamente con la frecuencia de rastreo de la página. Valores aceptados:   * always * hourly * daily * weekly * monthly * yearly * never   El valor "always" (siempre) debe utilizarse para describir documentos que cambian cada vez que se obtiene acceso a ellos. El valor "never" (nunca) debe utilizarse para describir direcciones URL archivadas.  Tenga en cuenta que el valor de esta etiqueta se considera una *sugerencia* y no una orden. A pesar de que los rastreadores de motores de búsqueda puedan tener en cuenta esta información a la hora de tomar decisiones, pueden rastrear páginas marcadas "hourly" (cada hora) con menor frecuencia de lo que indica la marca, así como rastrear páginas marcadas "yearly" (cada año) con más asiduidad. Asimismo, pueden rastrear periódicamente páginas marcadas "never" (nunca) para poder manejar los cambios inesperados que se produzcan en ellas. |
| <priority> | opcional | La prioridad de esta dirección URL es relativa con respecto a las demás URL de su sitio. Los valores válidos abarcan desde 0,0 a 1,0.  Este valor no afecta a la comparación de sus páginas con respecto a las de otros sitios; únicamente permite informar a los motores de búsqueda de las páginas que considera más importantes para los rastreadores.  La prioridad predeterminada de una página es 0,5.  Tenga en cuenta que la prioridad que asigne a la página no suele influir en la posición de sus URL en las páginas de resultados de los motores de búsqueda.  Los motores de búsqueda pueden utilizar esta información para elegir entre varias URL del mismo sitio, de modo que puede emplear esta etiqueta para incrementar las probabilidades de que sus páginas más importantes se incluyan en un índice de búsqueda.  Asimismo, tenga en cuenta que la asignación de alta prioridad a todas las URL de su sitio probablemente no le servirá de ayuda, dado que la prioridad es relativa y sólo se utiliza para elegir entre las distintas URL de su sitio. |

#### Caracteres de escape de entidad

El archivo de Sitemap debe tener codificación UTF-8; habitualmente puede establecerlo así al guardar el archivo. Al igual que con los archivos XML, los valores de datos (incluidas las URL) deben utilizar caracteres de escape de entidad para los caracteres de la tabla que encontrará más abajo.

|  |  |  |
| --- | --- | --- |
| **Carácter** | | **Código de caracteres de escape** |
| Símbolo de unión | & | &amp; |
| Comillas simples | ' | &apos; |
| Comillas | " | &quot; |
| Mayor que | > | &gt; |
| Menor que | < | &lt; |

#### Sitemap XML de ejemplo

El siguiente ejemplo muestra un Sitemap en formato XML. El Sitemap en cuestión contiene un número reducido de URL, cada una de las cuales presenta un conjunto diferente de parámetros opcionales.

<?xml version="1.0" encoding="UTF-8"?>

<[urlset](https://www.sitemaps.org/es/protocol.html#urlsetdef) xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">

<[url](https://www.sitemaps.org/es/protocol.html#urldef)>

<[loc](https://www.sitemaps.org/es/protocol.html#locdef)> [http://www.example.com/</loc](http://www.example.com/%3c/loc)>

<[lastmod](https://www.sitemaps.org/es/protocol.html#lastmoddef)>2005-01-01</lastmod>

<[changefreq](https://www.sitemaps.org/es/protocol.html#changefreqdef)>monthly</changefreq>

<[priority](https://www.sitemaps.org/es/protocol.html#prioritydef)>0.8</priority>

</url>

<[url](https://www.sitemaps.org/es/protocol.html#urldef)>

<[loc](https://www.sitemaps.org/es/protocol.html#locdef)>http://www.example.com/catalog?item=12&amp;desc=vacation\_hawaii</loc>

<[changefreq](https://www.sitemaps.org/es/protocol.html#changefreqdef)>weekly</changefreq>

   </url>

   <[url](https://www.sitemaps.org/es/protocol.html#urldef)>

      <[loc](https://www.sitemaps.org/es/protocol.html#locdef)>http://www.example.com/catalog?item=73&amp;desc=vacation\_new\_zealand</loc>

      <[lastmod](https://www.sitemaps.org/es/protocol.html#lastmoddef)>2004-12-23</lastmod>

      <[changefreq](https://www.sitemaps.org/es/protocol.html#changefreqdef)>weekly</changefreq>

   </url>

   <[url](https://www.sitemaps.org/es/protocol.html#urldef)>

      <[loc](https://www.sitemaps.org/es/protocol.html#locdef)>http://www.example.com/catalog?item=74&amp;desc=vacation\_newfoundland</loc>

      <[lastmod](https://www.sitemaps.org/es/protocol.html#lastmoddef)>2004-12-23T18:00:15+00:00</lastmod>

      <[priority](https://www.sitemaps.org/es/protocol.html#prioritydef)>0.3</priority>

   </url>

   <[url](https://www.sitemaps.org/es/protocol.html#urldef)>

      <[loc](https://www.sitemaps.org/es/protocol.html#locdef)>http://www.example.com/catalog?item=83&amp;desc=vacation\_usa</loc>

      <[lastmod](https://www.sitemaps.org/es/protocol.html#lastmoddef)>2004-11-23</lastmod>

   </url>

</urlset>

### Uso de archivos de índice de Sitemap (para agrupar varios archivos de Sitemap)

Puede proporcionar varios archivos de Sitemap, pero cada uno de ellos deberá contener un máximo de 50.000 direcciones URL y no superar los 50 MB (52,428,800 bytes), aunque se pueden comprimir

Si se quiere incluir más de 50.000 direcciones URL, deberá crear varios archivos de Sitemap.

Si proporciona varios Sitemaps, debe enumerarlos todos en un archivo de índice de Sitemap.

Los archivos de índice de Sitemap no pueden contener más de 50.000 Sitemaps y no deben superar los 50 MB (52,428,800 bytes), aunque se pueden comprimir.

Es posible disponer de más de un archivo de índice de Sitemap. El formato XML de un archivo de índice de Sitemap es muy parecido al formato XML de un archivo de Sitemap.

El archivo de índice de Sitemap debe:

* Comenzar con una etiqueta de apertura <[sitemapindex](https://www.sitemaps.org/es/protocol.html#sitemapIndex_sitemapindex)> y terminar con una de cierre </sitemapindex>.
* Incluir una entrada <[sitemap](https://www.sitemaps.org/es/protocol.html#sitemapIndex_sitemap)> para cada Sitemap como una etiqueta XML principal.
* Incluir una entrada secundaria <[loc](https://www.sitemaps.org/es/protocol.html#sitemapIndex_loc)> para cada etiqueta principal <sitemap>.

La etiqueta opcional <[lastmod](https://www.sitemaps.org/es/protocol.html#sitemapIndex_lastmod)> también está disponible para archivos de índice de Sitemap.

**Nota:** Un archivo de índice de Sitemap sólo puede especificar Sitemaps que se encuentren en la misma ubicación que el archivo de índice de Sitemap.

Por ejemplo, http://www.susitio.es/sitemap\_index.xml puede incluir Sitemaps en http://www.susitio.es, pero no en http://www.ejemplo.es o http://suhost.susitio.es. Igual que ocurre con los Sitemaps, el archivo de índice de su Sitemap debe estar codificado en UTF-8.

#### Índice XML de Sitemap de ejemplo

El siguiente ejemplo muestra un índice de Sitemap que incluye dos Sitemaps:

<?xml version="1.0" encoding="UTF-8"?>

<[sitemapindex](https://www.sitemaps.org/es/protocol.html#sitemapIndex_sitemapindex) xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">

   <[sitemap](https://www.sitemaps.org/es/protocol.html#sitemapIndex_sitemap)>

      <[loc](https://www.sitemaps.org/es/protocol.html#sitemapIndex_loc)>http://www.example.com/sitemap1.xml.gz</loc>

      <[lastmod](https://www.sitemaps.org/es/protocol.html#sitemapIndex_lastmod)>2004-10-01T18:23:17+00:00</lastmod>

   </sitemap>

   <[sitemap](https://www.sitemaps.org/es/protocol.html#sitemapIndex_sitemap)>

      <[loc](https://www.sitemaps.org/es/protocol.html#sitemapIndex_loc)>http://www.example.com/sitemap2.xml.gz</loc>

      <[lastmod](https://www.sitemaps.org/es/protocol.html#sitemapIndex_lastmod)>2005-01-01</lastmod>

   </sitemap>

</sitemapindex>

#### Definiciones de etiquetas XML de índice de Sitemap

|  |  |  |
| --- | --- | --- |
| **Atributo** |  | **Descripción** |
| <sitemapindex> | **obligatorio** | Encapsula información acerca de todos los Sitemaps del archivo. |
| <sitemap> | **obligatorio** | Encapsula información acerca de un Sitemap concreto. |
| <loc> | **obligatorio** | Identifica la ubicación del Sitemap.  Esta ubicación puede ser un Sitemap, un archivo Atom, un archivo RSS o un archivo de texto. |
| <lastmod> | opcional | Identifica la hora a la que se modificó el Sitemap correspondiente. No es la hora de modificación de alguna de las páginas incluidas en el Sitemap. El valor de la etiqueta lastmod debe encontrarse en formato [Fecha y hora de W3C](http://www.w3.org/TR/NOTE-datetime).  Al proporcionar la marca horaria de la última modificación, permite que los rastreadores de motores de búsqueda obtengan únicamente un subconjunto de los Sitemaps del índice, es decir, el rastreador sólo podrá obtener Sitemaps que han sido modificados a partir de una fecha concreta. Este mecanismo de obtención de Sitemap incremental permite un rápido descubrimiento de nuevas URL en sitios de gran tamaño. |

### Ubicación de archivos de Sitemap

* La ubicación de un archivo de Sitemap determina el grupo de URL que se pueden incluir en ese Sitemap.

Un archivo de Sitemap ubicado en http://ejemplo.es/catalog/sitemap.xml puede incluir URL que empiecen por http://ejemplo.es/catalog/, pero no URL que empiecen por <http://ejemplo.es/images/>.

* Si tiene permiso para cambiar http://ejemplo.org/ruta/sitemap.xml, puede suponer que también tiene permiso para proporcionar información para las URL con el prefijo <http://ejemplo.org/ruta/>.
* Todas las URL enumeradas en el Sitemap deben utilizar el mismo protocolo (http/https) y residir en el mismo host que el Sitemap.

Por ejemplo, si el Sitemap se encuentra en http://www.ejemplo.es/sitemap.xml,no puede incluir URL de [*http://subdominio*.ejemplo.es](http://subdominio.ejemplo.es) ni en [*https*://www.ejemplo.es/](https://www.ejemplo.es/)

* Las URL que no se estiman válidas dejan de tenerse en cuenta.
* Si utiliza una ruta con un número de puerto para enviar un Sitemap, debe incluir dicho número de puerto de la ruta en todas las URL que aparecen en el archivo del Sitemap. Por ejemplo, si el Sitemap se encuentra en http://www.ejemplo.es*:100*/sitemap.xml, todas las URL que aparecen en él deben empezar por <http://www.ejemplo.es:100>.

### Validación de Sitemap

Los siguientes esquemas XML definen los elementos y atributos que pueden aparecer en su archivo de Sitemap. Puede descargar este esquema desde cualquiera de los siguientes vínculos:

Existen varias herramientas que le pueden ayudar a validar la estructura de su Sitemap según este esquema. Puede encontrar una lista de herramientas útiles para XML en las siguientes ubicaciones:  
  
<http://www.w3.org/XML/Schema#Tools>  
<http://www.xml.com/pub/a/2000/12/13/schematools.html>

Para poder validar su archivo de Sitemap o su archivo de índice de Sitemap en función de un esquema, el archivo XML necesitará cabeceras adicionales, tal y como se muestra a continuación.

<?xml version='1.0' encoding='UTF-8'?>

<urlset xmlns:xsi=<http://www.w3.org/2001/XMLSchema-instance> xsi:schemaLocation=[http://www.sitemaps.org/schemas/sitemap/0.9 http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd](http://www.sitemaps.org/schemas/sitemap/0.9%20http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd) xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">

  <url>

     ...

  </url>

</urlset>

**Archivo de índice de Sitemap:**

<?xml version='1.0' encoding='UTF-8'?>

<sitemapindex xmlns:xsi=<http://www.w3.org/2001/XMLSchema-instance> xsi:schemaLocation=[http://www.sitemaps.org/schemas/sitemap/0.9 http://www.sitemaps.org/schemas/sitemap/0.9/siteindex.xsd](http://www.sitemaps.org/schemas/sitemap/0.9%20http://www.sitemaps.org/schemas/sitemap/0.9/siteindex.xsd) xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">

<sitemap>

      ...

   </sitemap>

</sitemapindex>