



## Sitemap, Mapa del sitio web. Importancia para el posicionamiento en buscadores

Fecha Domingo, 20 julio a las 02:10:51

Tema Internet y comunicaciones

Establecer un formato correcto para el mapa del sitio web (sitemap), es fundamental para que nuestra web sea escaneada completamente por los robots de búsqueda. Los motores de búsqueda (Google, Yahoo, MSN, etc) utilizan equipos rastreadores que visitan nuestra web y extraen su contenido, de esta forma se [mejora el posicionamiento](#) en los buscadores. Tener correctamente configurado el mapa del sitio web (sitemap) es fundamental para que el rastreo por parte de los robots sea lo más eficiente posible.

## Sitemap de texto plano (sin formato)

El Sitemap (mapa del sitio web) es un fichero de texto plano (aunque pueden haber otros formatos, como XML) donde se deben guardar todas las URLs (direcciones) de nuestro sitio, a las que queramos que accedan los motores de búsqueda. Cada línea de este fichero contendrá una URL completa, por ejemplo:

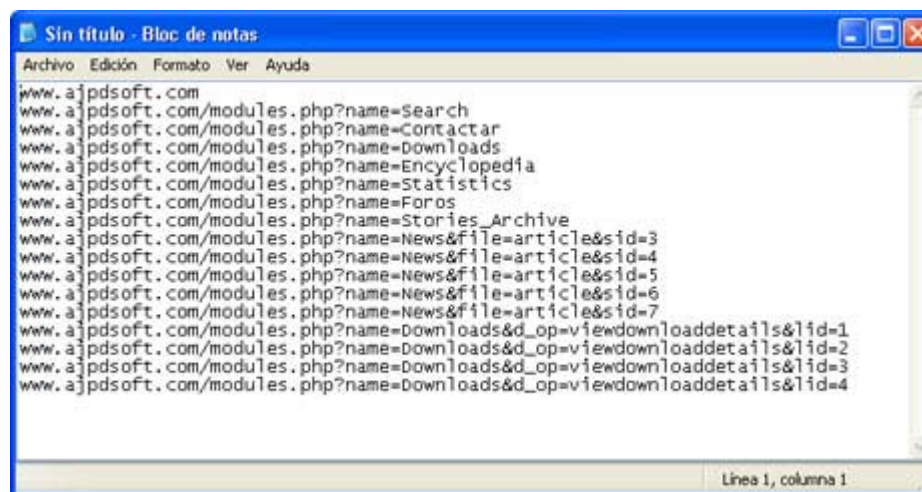
```
www.ajpdsoft.com/modules.php?name=News&file=article&sid=297
www.ajpdsoft.com/modules.php?name=Quienes_Somos
www.ajpdsoft.com/modules.php?name=Downloads
www.ajpdsoft.com/modules.php?name=Stories_Archive&sa=show_all
```

### Directrices a seguir para crear un fichero de sitemap de texto plano

- Las direcciones (URL) deben ser **completas**, puesto que los robots de los motores de búsqueda rastrean las direcciones exactamente como se indiquen.
- Para el caso de Google, cada fichero sitemap puede contener un **máximo** de 50.000 URL. Si es necesario superar esta cifra, los ficheros sitemap se pueden dividir en varios.
- La **codificación** (juego de caracteres) del archivo de texto debe ser UTF-8.
- El archivo de texto **sólo** debe contener URLs, sin ningún tipo de encabezado ni pie.
- Como **nombre** de este archivo, se puede utilizar el que se quiera. Google recomienda que la extensión del archivo sea .txt, de modo que se pueda identificar como archivo de texto, por ejemplo, sitemap.txt.

### Creación del fichero sitemap de texto plano

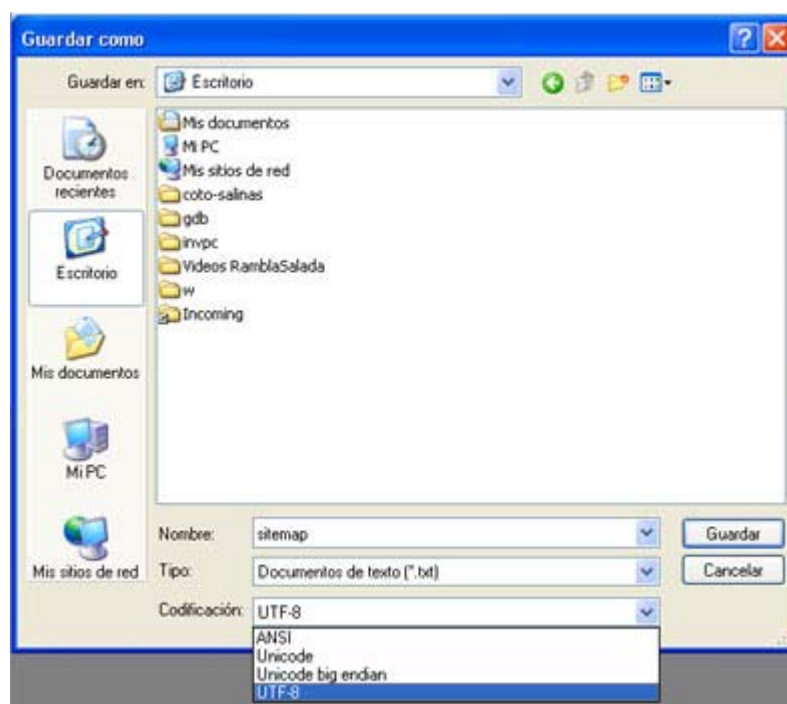
Para crear este fichero abriremos cualquier editor de textos sin formato, por ejemplo *Notepad* (de Windows) o *vi* (de Linux). En nuestro caso utilizaremos Notepad (Bloc de notas) de Windows. En el editor introduciremos una línea por cada URL con la dirección completa:



Tras añadir todas las URL de nuestro sitio web a las que queramos que accedan los robots de búsqueda, pulsaremos en "Archivo" - "Guardar como...":

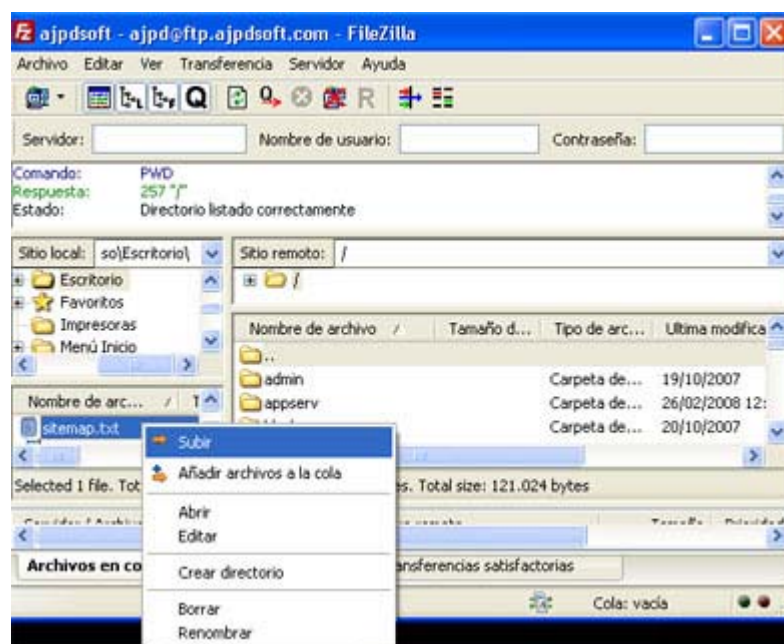


Seleccionaremos la carpeta donde guardaremos el fichero sitemap, introduciremos el nombre del fichero sitemap, por ejemplo "sitemap" y en el desplegable de "Codificación" seleccionaremos "UTF-8". Pulsaremos "Guardar":



Una vez creado el fichero de sitemap, con todas las URL de nuestro sitio web, será necesario, obviamente, subirlo al sitio web mediante FTP, para ello abriremos cualquier cliente de FTP,

nos conectaremos al sitio FTP y subiremos el fichero sitemap creado en el raíz del sitio:



## Cómo indicar a los robots de búsqueda (rastreadores) cómo localizar nuestro fichero sitemap

Para establecer el fichero de sitemap que hemos creado anteriormente (sitemap.txt), deberemos crear y colocar un fichero de texto plano en el raíz de nuestro sitio web llamado **robots.txt**. Este fichero se utiliza para varias cosas, algunas de ellas:

- Se utiliza para indicar a ciertos "bots" (equipos que acceden a nuestra web y obtienen el contenido para uso de buscadores y demás) a qué partes de nuestra web pueden acceder y a cuales no.
- También se utiliza para indicar la ubicación del fichero sitemap (mapa del sitio web).
- Se puede indicar a qué robots permitiremos el acceso y a cuales no.
- Indicar qué ficheros se indexarán, por ejemplo, se le puede indicar que indexe todos los ficheros salvo las imágenes:
- Indicar a los robots que indexen una página pero no los enlaces de esta página a otras:
- Para indicar qué robots (MSN, Yahoo, Google, etc.) puedan indexar una página y cuales no. Por ejemplo, podremos indicar que todos los robots salvo los de Google indexen una determinada página web:

Por ejemplo, si creamos el fichero robots.txt y con este contenido:

```
User-agent: *
Disallow:
```

Estaremos indicando que cualquier robots puede entrar y obtener el contenido de todos los ficheros del raíz de nuestro sitio web.

En este otro ejemplo indicamos que los robots no puedan acceder a ningún fichero de la raíz de nuestro sitio web:

```
User-agent: *
```

*Disallow: /*

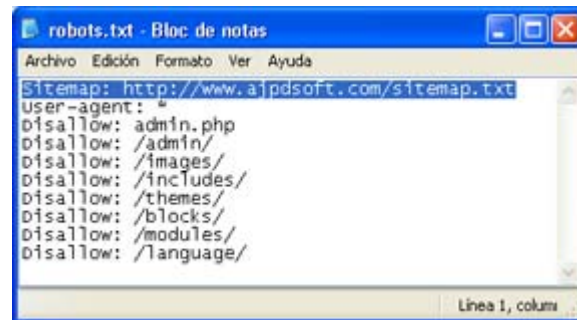
Volviendo al tema que nos ocupa, para indicar en este fichero la ubicación y nombre de nuestro fichero *sitemap.txt* añadiremos la siguiente línea:

*Sitemap:*

Donde será la URL completa del fichero *sitemap.txt*, por ejemplo, para nuestro caso, añadiremos esta línea al fichero *robots.txt*:

*Sitemap: http://www.ajpdsoft.com/sitemap.txt*

Un ejemplo de contenido de fichero *robots.txt*:



Guardaremos el fichero *robots.txt* y lo subiremos en el mismo sitio que el fichero *sitemap.txt* mediante un cliente FTP.

Con estos pasos, la próxima vez que un robots de cualquier motor de búsqueda (Yahoo, MSN, Google, etc) visite nuestro sitio web utilizará el fichero *sitemap.txt* como guía para indexar nuestro sitio web.

## Sitemap en formato XML

Otra posibilidad de formato del fichero *sitemap* puede ser en XML, siguiendo las siguientes pautas:

- En el encabezado del fichero colocaremos el siguiente texto:

```
< ?xml version="1.0" encoding="UTF-8"? >
< urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9" >
```

- Por cada URL añadiremos la siguiente etiqueta XML:

```
URL_Completa/
aaa-mm-dd
monthly
0.8
```

Donde:

- **URL\_Completa:** será la dirección URL de la página a indexar, por ejemplo: *http://www.ajpdsoft.com/modules.php?name=Foros*. Esta etiqueta es obligatoria.
- **lastmod:** será la fecha con formato año-mes-día de última modificación de la

URL. Esta etiqueta es opcional.

- **changefreq**: aquí indicaremos la frecuencia con la que cambiaremos esta página, de esta forma el robot sólo pasará cuando sea necesario y evitaremos tráfico innecesario en nuestra web. Los posibles valores para esta propiedad son: always, hourly, daily, weekly, monthly, yearly, never. Esta etiqueta es opcional.
  - **priority**: prioridad que se le dará a la indexación de esta URL. La prioridad va de 0.0 a 1.0, siendo 0.5 el valor por defecto. Esta etiqueta es opcional.
- Al final del fichero XML colocaremos la siguiente línea:

Un ejemplo de fichero sitemap en formato XML podría ser:

```

sitemap.txt - Bloc de notas
Archivo Edición Formato Ver Ayuda
<?xml version="1.0" encoding="UTF-8"?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
  <url>
    <loc>http://www.ajpdsoft.com/modules.php?name=search</loc>
    <lastmod>2008-07-20</lastmod>
    <changefreq>yearly</changefreq>
    <priority>0.1</priority>
  </url>
  <url>
    <loc>http://www.ajpdsoft.com/modules.php?name=Contactar</loc>
    <lastmod>2008-07-20</lastmod>
    <changefreq>yearly</changefreq>
    <priority>0.1</priority>
  </url>
  <url>
    <loc>http://www.ajpdsoft.com/modules.php?name=Downloads</loc>
    <lastmod>2008-07-20</lastmod>
    <changefreq>yearly</changefreq>
    <priority>0.5</priority>
  </url>
</urlset>
Línea 25, columna 1

```

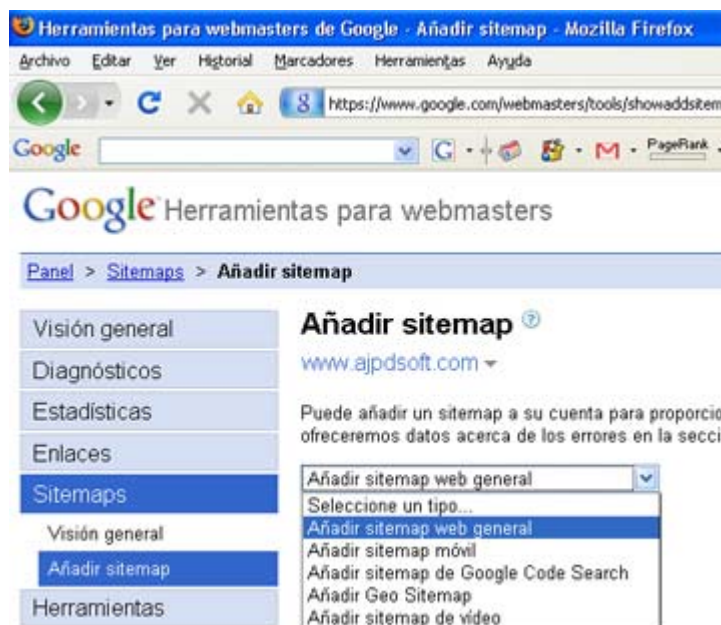
Al igual que para el fichero sitemap en formato de texto plano (.txt), este fichero deberemos guardarlo con codificación UTF-8, con extensión xml, por ejemplo sitemap.xml y subirlo a nuestro sitio web mediante un cliente de FTP. También deberemos indicar en el fichero robots.txt la ubicación y nombre de este fichero.

## Subir a Google el fichero sitemap directamente (mediante Google para Webmasters)

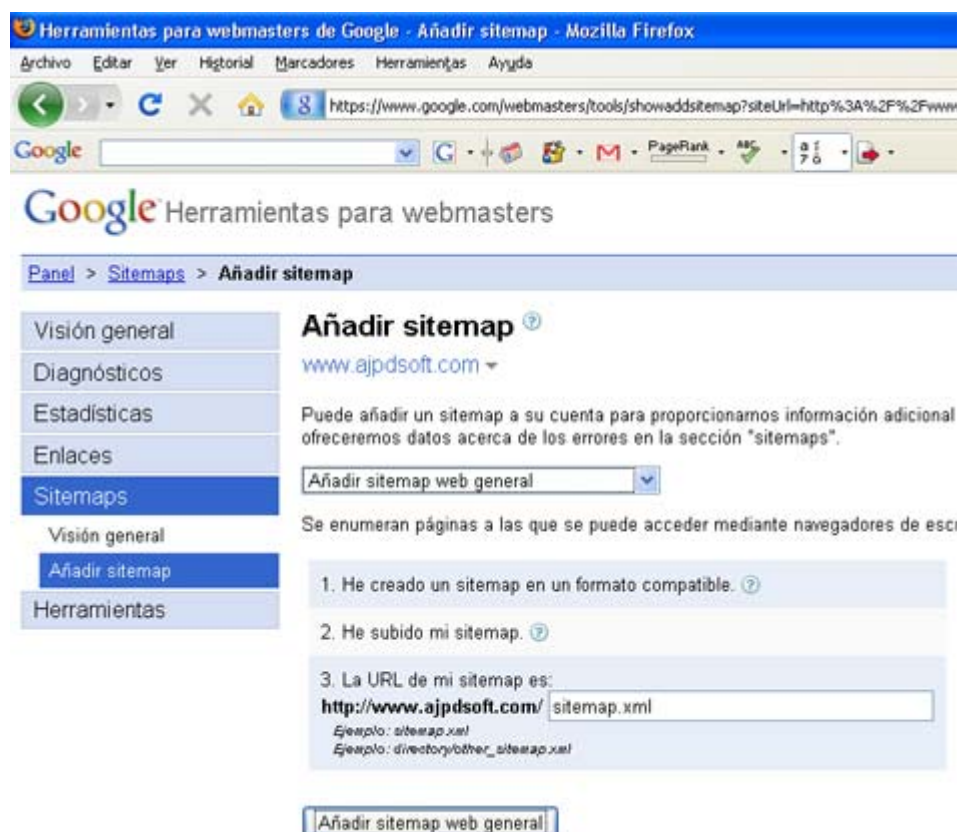
En primer lugar deberemos tener una cuenta en Google para Webmasters (gratuito), a continuación os mostramos la URL de un artículo sobre cómo funciona Google para Webmasters y las utilidades que incorpora:

[Google para Webmasters](#)

Desde la opción "Sitemaps" - "Añadir sitemap" seleccionaremos, en el desplegable "Añadir sitemap web general":



Nos aparecerá un cuadro de texto para que introduzcamos la ubicación y nombre de nuestro fichero sitemap, en nuestro caso `http://www.ajpdssoft.com/sitemap.xml`, tras introducir el nombre del fichero sitemap pulsaremos en "Añadir sitemap web general":



Lógicamente, para que este procedimiento funcione y Google pueda obtener el contenido del fichero sitemap debe estar subido a nuestro sitio web.

Si Google lo encuentra mostrará una ventana como esta:

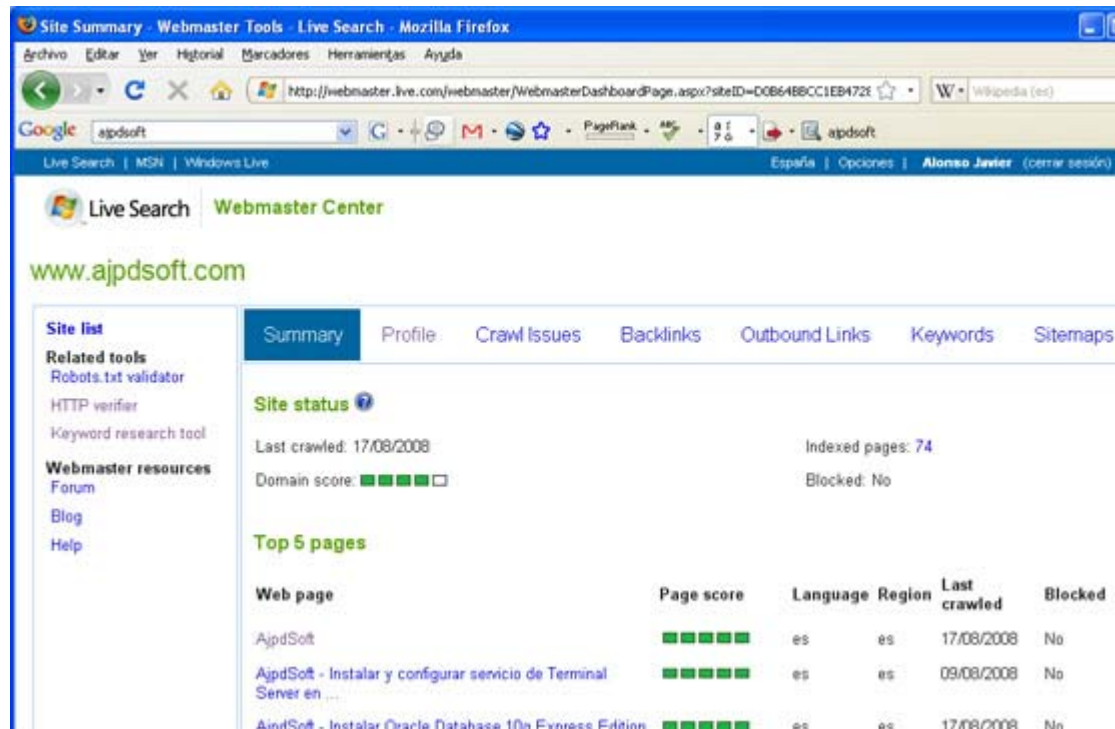


Con el texto: "Ha añadido <http://www.ajpdsoft.com/sitemap.xml>. Los informes pueden tardar varias horas en actualizarse. Gracias por su paciencia."

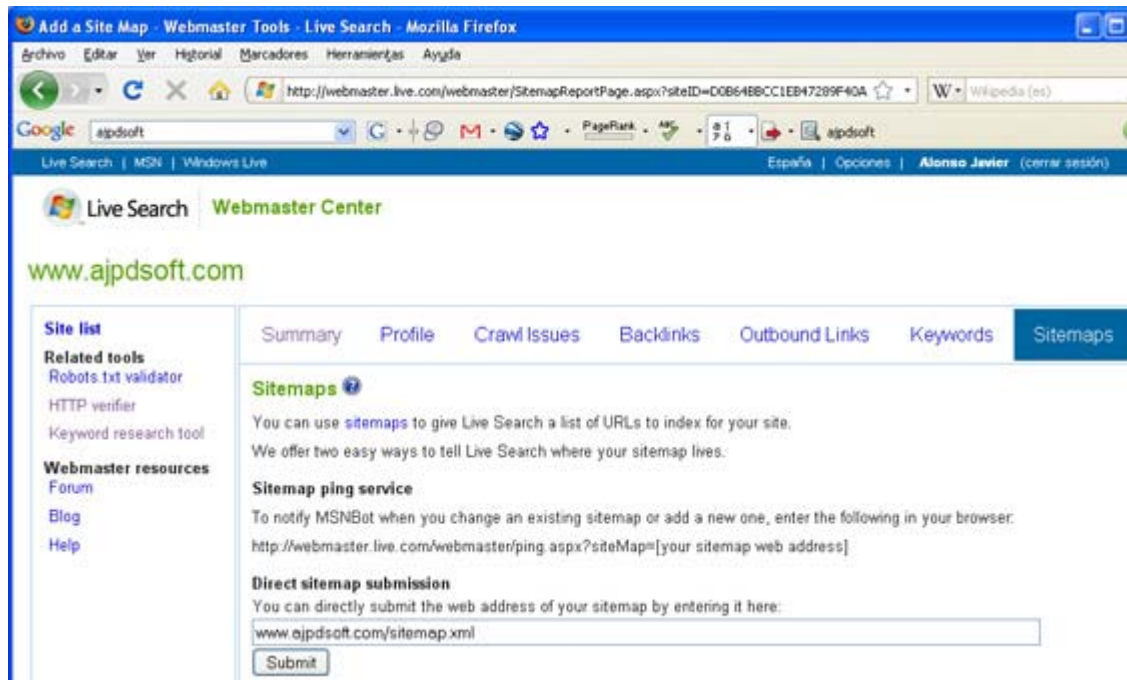
## Subir a MSN Live Search el fichero sitemap directamente (mediante Webmaster Center)

La forma de proceder es similar a Google para Webmasters, la URL para acceso a MSN Live Search Webmaster Center es: <http://webmaster.live.com>, necesitaremos disponer de una cuenta de Hotmail (al igual que en el caso de Google para Webmasters, que se necesita una de GMail). Tras realizar la validación de la web, subiendo un fichero .xml, con el contenido que indica el asistente de Webmaster Center, la web quedará validada. Podremos indicar la ubicación del fichero sitemap creado. Todo muy parecido a Google para Webmasters.

La sección "Summary" de **Live Search Webmaster Center** tiene este aspecto:

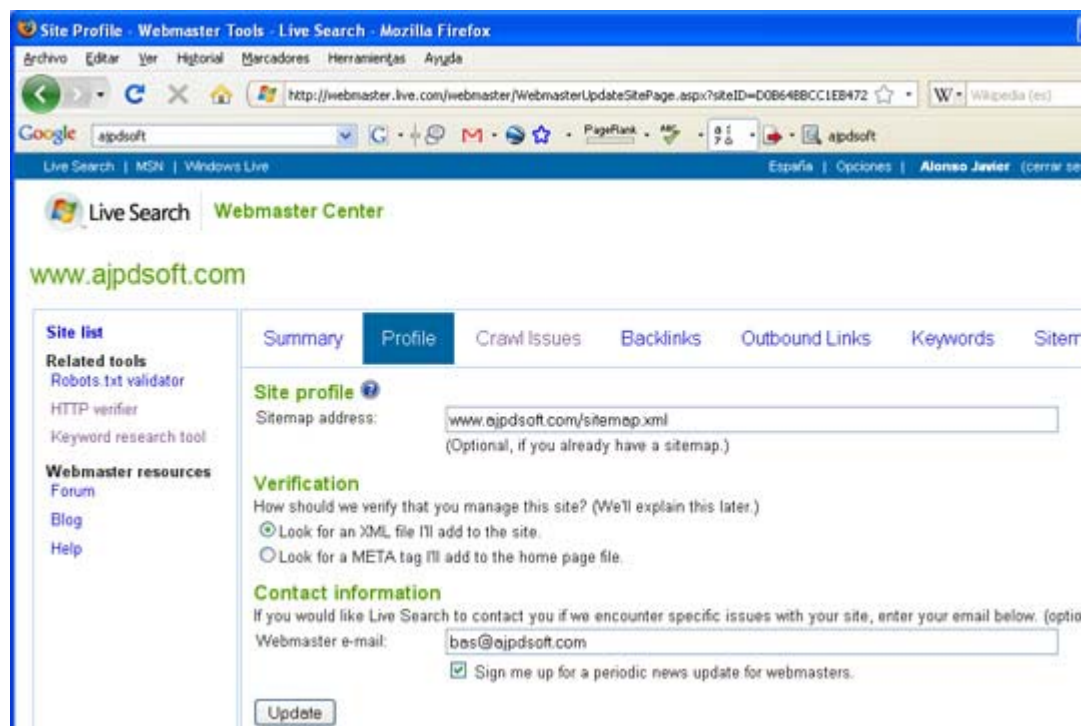


La sección "Sitemaps" de **Live Search Webmaster Center** podremos indicar la URL de nuestro fichero de sitemap:



También podremos hacerlo desde la sección "Profile":





## Conclusión

El fichero sitemap debe existir y estar actualizado en todo sitio web que se precie. Actualmente, salvo que gastemos una cantidad importante en publicidad, una web "no existe" si los motores de búsqueda no la han indexado correctamente. Cuanta más información de nuestra web contengan los motores de búsqueda más posibilidades de "aparecer" bien posicionados tendremos.

Siempre es recomendable utilizar el protocolo de sitemaps, ya que el mismo archivo puede enviarse a varios motores de búsqueda como Google, MSN y Yahoo, que son miembros de sitemaps.org. El formato idóneo de sitemaps es el XML, pues como se ha podido comprobar contiene más información para los robots, en cambio, el formato de texto plano (.txt) tan sólo aporta la URL.

Independientemente de que se haya subido el fichero de sitemaps al sitio web, siempre es recomendable utilizar las herramientas para webmasters de los diferentes motores de búsqueda para subir este fichero directamente al motor. Por ejemplo, para el caso de Google, se puede utilizar [Google para Webmasters](#) para subir este fichero. De esta forma no será necesario esperar a que los motores pasen por nuestro sitio para obtener el fichero sitemap, pues le estamos indicando directamente qué fichero utilizar y qué contenido tiene.

**Nota importante:** En el caso de los ficheros de sitemap con formato XML, si alguna de las URL incluye símbolos como "&" deberán ser reemplazados por su correspondiente carácter de escape: "&", de lo contrario es posible que la lectura del fichero sitemap.xml dé error.

Revisado por AjpdsSoft el 20-08-2008.

Este artículo proviene de AjpdSoft  
<http://www.ajpdsoft.com>

La dirección de esta noticia es:  
<http://www.ajpdsoft.com/modules.php?name=News&file=article&sid=338>