

Tema: Por qué son importantes el sitemap.xml y robots.txt para un sitio web

1. Imaginemos que nuestro sitio web es una BIBLIOTECA GIGANTE

Para que entiendan mejor:

- **El sitio web** = Una biblioteca enorme con miles de libros
- **Los buscadores (Google, etc.)** = Bibliotecarios que quieren organizar todos los libros
- **Las páginas web** = Los libros individuales en la biblioteca

PARTE 1: ¿Qué es y para qué sirve el robots.txt?


Paso 1: El problema

Imaginen que el bibliotecario llega a la biblioteca y:

- No sabe por dónde empezar
- No sabe qué secciones puede visitar
- No sabe qué libros están restringidos

Paso 2: La solución - El cartel de BIENVENIDA/INSTRUCCIONES

El robots.txt es como **un cartel en la entrada de la biblioteca** que dice:

 **"Hola bibliotecario, aquí tienes las reglas:"**

- "Puedes entrar en estas secciones: /libros/, /noticias/"
- "NO entres aquí: /archivos-privados/, /admin/"
- "Aquí tienes el mapa: /sitemap.xml"

Paso 3: Ejemplo real

robots.txt

User-agent: *

Allow: /

Disallow: /admin/

Disallow: /tmp/

Sitemap: <https://misitio.com/sitemap.xml>

Paso 4: ¿Por qué es importante?

- **Protege contenido privado:** Evita que los buscadores indexen áreas administrativas
- **Ahorra recursos:** Guía a los bots para que no pierdan tiempo en páginas irrelevantes
- **Control:** Tú decides qué se muestra en los buscadores

PARTE 2: ¿Qué es y para qué sirve el sitemap.xml?

Paso 1: El problema

El bibliotecario está en la biblioteca, pero:

- Hay libros escondidos en armarios
- Libros nuevos que no están en el catálogo
- No sabe qué libros son más importantes

Paso 2: La solución - El CATÁLOGO COMPLETO

El sitemap.xml es como **el catálogo oficial de la biblioteca** que contiene:

 **"Lista completa de todos los libros (páginas)":**

- Dirección exacta de cada página
- Cuándo se actualizó por última vez
- Qué tan importante es cada página
- Con qué frecuencia cambia

Paso 3: Ejemplo real

xml

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

```
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
```

```
<url>
```

```
<loc>https://misitio.com/</loc>
```

```
<lastmod>2024-01-15</lastmod>
```

```
<changefreq>monthly</changefreq>
```

```
<priority>1.0</priority>
```

```
</url>
```

```
<url>
```

```
<loc>https://misitio.com/nosotros</loc>
```

```
<lastmod>2024-01-10</lastmod>
```

<changefreq>yearly</changefreq>

<priority>0.8</priority>

</url>

</urlset>

Paso 4: ¿Por qué es importante?

- **Descubrimiento rápido:** Los buscadores encuentran todas tus páginas inmediatamente
- **Indexación eficiente:** Saben qué páginas son más importantes y con qué frecuencia actualizarlas
- **Contenido nuevo:** Notificas inmediatamente sobre nuevas páginas o actualizaciones

PARTE 3: RESUMEN

Robots.txt = EL GUARDIÁN

- **Función:** Dar instrucciones a los buscadores
- **Como si fuera:** El portero que dice "por aquí sí, por aquí no"
- **¿Por qué lo necesitas?:** Para proteger y organizar

Sitemap.xml = EL MAPA DEL TESORO

- **Función:** Mostrar TODAS las páginas importantes
- **Como si fuera:** El mapa que marca dónde está cada tesoro
- **¿Por qué lo necesitas?:** Para ser encontrado completamente y rápido

¿Qué pasa si NO los tienes?

- Los buscadores pueden perderse páginas importantes
- Pueden indexar contenido que no quieres que aparezca
- Tardarán más en descubrir tu contenido nuevo
- Perderás visibilidad en los resultados de búsqueda

Datos curiosos para captar atención

- El **95%** de los sitios web profesionales tienen ambos archivos
- **Google recomienda** siempre tener un sitemap.xml
- Sin robots.txt, los buscadores pueden indexar hasta tus archivos de configuración

Generar sitemap con: [XML Generator](#)