

Módulo 1. Factores de estructura on page

Unidad 1. Factores de estructura on page.

El *Search Engine Optimization* o posicionamiento en buscadores (SEO por sus siglas en inglés) es una disciplina bastante amplia y, por tal motivo, es necesario tomar en consideración una gran cantidad de elementos que influyen en la optimización de los resultados obtenidos en los motores de búsqueda. La verdadera razón de querer hacer SEO es la capacidad de conseguir tráfico de calidad al lograr entender la intencionalidad del usuario y proporcionar contenido seductor para ese momento. Esto elevará la probabilidad de ser elegido entre millones de resultados orgánicos posibles y, mientras más certero sea en este aspecto, mejor será el resultado.

Todo esto es muy importante, pero, si se lo deja solo en decir intencionalidad o contenido seductor, queda en sentido muy amplio y ambiguo. En la medida en que se sea más específico con lo que hay que mejorar y que se entienda más precisamente lo que están buscando los algoritmos de búsqueda como el de Google, se podrá generar un diagnóstico lo más cercano posible a lo que los buscadores considerarán relevante y, por ende, trabajar en mejorar con el fin de que esta situación genere sus frutos.

Si se analizan los resultados orgánicos en los motores de búsqueda, se podrá apreciar que existen elementos sumamente importantes que no se pueden pasar por alto, por ejemplo:

Figura 1. Ejemplo de un resultado de búsqueda orgánica en Google



Fuente: Adaptación propia con base en captura parcial de pantalla del buscador de Google.

Este módulo está dedicado a entender los factores que, dentro de la estructura *on page*, pueden ayudar a generar los resultados esperados. A continuación, se presenta un listado de estos factores por categoría.

1.1.1 Importancia de títulos y metatítulos

Un título es la frase principal que describe un contenido y a partir del cual muchas veces se decide si es interesante leerlo o no. Un metatítulo es la expresión de dicho título a través de una metaetiqueta en el HTML (*hypertext markup language*) de cada página web. Para desarrollarse en estos aspectos, se van a estudiar los factores de optimización:

- Factor 1: uso de etiquetas *title*. En el HTML de los sitios web, existe una interfaz HTML que presenta el contenido y estos aplican condiciones o presentan información que después es tomada en cuenta para los motores de búsqueda y la evaluación que hacen de cada caso. La etiqueta *title* es necesaria para declarar el título de la página a través del uso de la siguiente estructura de etiqueta:

```
<title>Título de la página</title>
```

Este título va colocado en la sección *head* del HTML y, normalmente, suele ser una de las primeras instrucciones presentadas en el sitio web. Usar estas etiquetas en todas las páginas de su sitio web debe ser una de las mayores prioridades en cuestiones de SEO.

- Factor 2: títulos únicos por página. Los títulos mencionados en el factor anterior deberán de ser únicos y no repetirse en ninguna página dentro del sitio web. La idea central de este punto es darle originalidad a cada página y que el contenido expresado sea distinto entre sí. Esto ayudará a

diferenciar para qué temas se tomará una página u otra como relevante y se descartarán frases genéricas que, simplemente, se repliquen en todo el sitio.

- Factor 3: información relevante en *title*: este factor consta de la premisa de que los títulos deben contener las palabras clave más relevantes para lo que se está hablando. Esto quiere decir que, al utilizar palabras más buscadas o con mayor potencial de posicionamiento, se podrá alcanzar a una mayor audiencia y, gracias a ello, conseguir mayor y mejor tráfico de calidad.
- Factor 4: nombre de empresa, sitio o ubicación relevante. Esto está enfocado a que, como parte del título, se coloque el nombre de la empresa a la que pertenece el sitio web o de la ubicación en la que se desea destacar; por ejemplo, se podría aplicar una estructura como la siguiente: "{nombre de la empresa}: La mayor calidad en {nombre de producto} en todo {nombre de región, país o ciudad}".
- Factor 5: respetar la longitud máxima en título. La idea de respetar una longitud particular en los títulos es para que, al aparecer en los resultados de búsqueda el elemento en cuestión, no exceda la cantidad de espacio disponible y que el concepto que expresa quede cortado o con puntos suspensivos. Esto reduciría el nivel de clickeo que los usuarios hagan sobre el resultado al mismo tiempo que se tomaría como menos relevante al no respetar las normas básicas de SEO. La longitud máxima en Google para un título es de 70 caracteres (incluyendo espacios), aunque se recomienda que, en la medida de lo posible, no exceda de 65 caracteres para poder aprovechar la utilización de otros elementos como microformatos y de elementos enriquecidos (los cuales se verán más adelante en este curso).
- Factor 6: respetar la longitud mínima en título. De la misma forma y con el mismo sentido que se recomienda respetar la longitud máxima, también hay recomendaciones para la longitud mínima. Si bien no existe un compendio técnico que defina cuál es la extensión mínima de un título, es muy común que los títulos menores a 30 caracteres (incluyendo espacios) tengan menor desempeño en los resultados orgánicos debido a que, en su corta extensión, hay muy poco que se puede relatar sobre el contenido que posee la página de destino y, por eso mismo, se está desaprovechando el uso de palabras clave que podrían aumentar su alcance.
- Factor 7: evitar palabras clave innecesarias en el título. De la misma manera en que, anteriormente, se destacaba el uso de palabras clave en los títulos, también es importante destacar que la redacción debe ser natural y apropiada para que tenga sentido ante la lectura. Esto quiere decir que hay que evitar todo tipo de iniciativa en donde el título pueda ser solo una sumatoria de palabras clave sin un sentido narrativo conciso. Además, cuando se enfoca en muchas temáticas a la vez, es más complicado ser relevante que cuando se toma un tema o una palabra clave bien específicos para titular el contenido.

1.1.2 Importancia de las metadescripciones

Las metadescripciones, en esencia, no son un factor de importancia SEO en el sentido de que la inclusión o no de palabras clave no afecta en ningún modo el ranking de un sitio web. No obstante, se tratan de factores primordiales que, en su correcto uso, potencian dramáticamente el porcentaje de clic que un sitio puede recibir desde los buscadores. Esto se debe a que los buscadores usan esta etiqueta como la porción de texto descriptiva (*snippet*) en los resultados de búsqueda. Con ello, se pretende mejorar el CTR (relación entre clics e impresiones) de una nota, ya que ser clicado más veces sí influiría directamente en el posicionamiento en el buscador.

Una buena metadescripción debe tener un uso y una distribución inteligente de las palabras clave, debe incluir un texto atractivo que incite al usuario a visitar el sitio y, además, debe adecuarse a los límites de caracteres que tienen los buscadores (una longitud óptima es entre 145 y 155 caracteres).

Figura 2: Snippet de resultado de búsqueda con buena descripción



Fuente: Adaptación propia con base en captura parcial de pantalla del buscador de Google.

También es importante resaltar que Google pone en negrita el contenido del título o metadescripción que concuerda con la consulta del usuario, lo que aumenta de este modo el impacto visual y, por ende, las chances de que el usuario haga clic en el sitio.

Figura 3. Snippet de resultado de búsqueda con mala descripción



Fuente: Adaptación propia con base en captura parcial de pantalla del buscador de Google.

Cuando la metadescripción es inexistente o demasiado breve, Google extrae el *snippet* desde el contenido principal de la página, lo que vuelve irrelevante la metadescripción y se asume que genera poco valor para el usuario.

Para saber cómo generar un buen uso de las metadescripciones, se sugiere considerar los siguientes factores:

- Factor 8: uso de la metaetiqueta *description* en todas las páginas del sitio. Al igual que el uso de la etiqueta de títulos, la etiqueta de metadescripciones debe utilizarse en todas las páginas del sitio web con la siguiente estructura en el *head*:

```
<meta name="description" content="Texto de la meta descripción" />
```

- Factor 9: descripción independiente para cada página. De la misma forma que sucede con los títulos, las metadescripciones también deben ser únicas y originales entre todas las páginas del sitio web. Esto brindará originalidad y especificidad en lo que plantea el contenido y generará un elemento único y específico, el cual será más fácil de posicionar según la intencionalidad del usuario a la hora de seleccionar el resultado más apropiado para su búsqueda.
- Factor 10: duplicidad de información contra el *title*. Nunca debería duplicarse la información contenida en títulos y descripciones. Se recomienda que se utilice información que pueda complementarse y que se trabaje en el proceso de convencimiento del usuario de que este resultado es el más apropiado para su búsqueda.
- Factor 11: inclusión de *call to action* y redacción en voz activa y accionable. La metadescripción debe estar redactada en forma natural y fácil de comprender para el que navega los resultados de búsqueda, pero, al mismo tiempo, debe ser sugestiva para apoyar la decisión de hacer clic sobre él. Redactar la descripción en una voz activa que fomente tomar una acción es uno de los elementos que mejor funcionan a la hora de optimizar este elemento para conseguir mejores tasas de cliqueo.

- Factor 12: coincidencia con el contenido. Se recomienda que, para que la metadescripción sea verdaderamente efectiva, se incluya el mismo contenido al que se hace referencia en el primer párrafo del sitio. Esto permitirá generar una mejor experiencia de usuario, así como un mayor impacto en quien ingresa al sitio y se pretende que ejecute una conversión.

1.1.3 Estructura de URL con sus dominios y subdominios

Las *uniform resource locator* (URL por sus siglas en inglés), o localizador uniforme de recursos, están compuestas por cuatro factores importantes: los subdominios, el dominio, la extensión de dominio y la URI (*uniform resource identifier*). Estos factores son parte fundamental en cualquier URL y están presentes para determinar el tipo de contenido y la ubicación a la cual se está accediendo.

Figura 4. Partes de una URL



Fuente: Elaboración propia.

Por ejemplo, en el siguiente caso de URL:

finance.yahoo.com/quote/%5EMERV?p=%5EMERV

El subdominio en esta URL es la palabra *finance*. El dominio es Yahoo con su extensión de dominio *.com*. A partir de allí, comienza la URI, la cual constituye todo lo que está después de la diagonal posterior a la extensión.

Existe un número de factores a nivel estructural que puede afectar negativamente el rendimiento del sitio en buscadores como la jerarquía de directorios, inclusión de ID de sesión, una profundidad de directorios superior a tres o poseer una

extensión superior a los 65 caracteres. Muchos de estos puntos son cruciales para que la indexación sea completa y para que el resultado de búsqueda pueda ser amigable con quien está evaluando la selección. De esta manera se consigue su poder sobre el ecosistema de SEO y es aquí donde se atribuirá la relevancia que se detecte sobre cualquier página. Todo esto se le atribuye a la URL, por lo que, con una responsabilidad tan grande, lo menos que se puede hacer es optimizarla para que, en ella misma, también se puedan sumar puntos de relevancia.

Para aprovechar el potencial que la URL presenta, es necesario cumplir con los siguientes factores:

- Factor 13: manejo de categorías descriptivas. La idea de optimización para cada URL comienza desde el concepto de *URL amigable*. Este concepto hace referencia a que las URL sean legibles y significativas para la interpretación humana y que, al mismo tiempo, los niveles de navegación que representa tengan sentido. Por ejemplo, una URL no amigable se vería de la siguiente manera:

www.dominio.com/lkasdj2342/23/2342lkj/ueucc345345id=45678?tmecsa=3234

En cambio, esa misma URL, si fuera amigable, podría verse de la siguiente manera:

www.dominio.com/paquetes-de-viaje/rio-de-janeiro.

Siempre deben mantenerse URL amigables para que sea más fácil entender el contenido que se almacena en dicha ubicación.

- Factor 14: manejo de nombres de archivo para los documentos del sitio. En el caso de que la parte final de una URL haga referencia a la ubicación de un documento (por ejemplo: un PDF) se recomienda que esta haga referencia específica y amigable al nombre del archivo y a una extensión estándar de su formato. Esto ayudará a indexarlo más fácilmente. Por ejemplo: www.dominio.com/paquetes-de-viaje/rio-de-janeiro/itinerario-de-paseos.pdf
- Factor 15: manejo de URL estática contra dinámica. Una URL estática es la que no va mutando o cambiando según algún parámetro, sino que siempre se mantiene igual. Por otro lado, la URL dinámica es aquella que, según distintos filtros, segmentos o tipos de navegación (entre otros posibles elementos) va cambiando constantemente y el contenido todavía se mantiene igual en gran medida. Las URL dinámicas generan un gran problema de contenido duplicado en los motores de búsqueda y son dañinos para la

relevancia de SEO, principalmente, debido a la atomización de la referencia de calidad que se podría transmitir.

Por ejemplo, una URL que cambia o agrega parámetros según filtros, como en el caso de un sitio web turístico, en el que se genera una búsqueda de un vuelo entre dos destinos para una fecha en particular. Esto genera una gran cantidad de parámetros que, si se cambia la fecha, estos parámetros también cambiarán, aunque se muestre el mismo contenido. Esto genera duplicación de contenido y es contraproducente a nivel de SEO.

El ejemplo descrito arriba se podría ver reflejado en estas dos URL de ejemplo:

URL con fecha 1: http://www.avantrip.com/vuelos/rio-de-janeiro-desde-buenos-aires?av-seleccion-grupo=on&isMulticity=false&destinationFromId%5B0%5D=BUE&round_trip=&destinationToId%5B0%5D=RIO&dateFrom%5B0%5D=10-03-2017&dateTo=24-03-2017&adults=1&children=0&babies=0&flightClass=NMO.GBL.SCL.ECO&hostname=www.avantrip.com

URL con fecha 2: http://www.avantrip.com/vuelos/rio-de-janeiro-desde-buenos-aires?av-seleccion-grupo=on&isMulticity=false&destinationFromId%5B0%5D=BUE&round_trip=&destinationToId%5B0%5D=RIO&dateFrom%5B0%5D=17-03-2017&dateTo=31-03-2017&adults=1&children=0&babies=0&flightClass=NMO.GBL.SCL.ECO&hostname=www.avantrip.com

Pese a que lo único que cambió fue la fecha y el contenido y el resultado fue el mismo, la URL es diferente, con lo cual se da una duplicación de contenidos. Lo ideal es que las URL dinámicas puedan respetar una estructura estática.

- Factor 16: utilización de signos de puntuación entre palabras de la URL. Las tildes, comas, dos puntos, punto y coma, etcétera no son parámetros aceptados en las URL. Por lo tanto, son elementos no amigables con la optimización SEO. El uso de estos perjudicará el posicionamiento orgánico debido a que se reconocen como elementos no estándar (caracteres Non-ASCII) que no brindan información relevante acorde a la estructura indicada por las buenas prácticas.
- Factor 17: abstinencia de parámetros irrelevantes en la URL. En la misma línea de los dos factores anteriores, es posible entender que existen URL que incluyen en su estructura algunos parámetros que lo único que hacen es dificultar la lectura humana al complicar su estructura y evitar una visualización limpia de lo que se busca y los niveles a los que pertenece el contenido.

Utilizando el mismo ejemplo del factor 15, se pueden apreciar los siguientes parámetros irrelevantes a nivel SEO:

http://www.avantrip.com/vuelos/rio-de-janeiro-desde-buenos-aires?av-seleccion-grupo=on&isMulticity=false&destinationFromId%5B0%5D=BUE&round_trip=&destinationToId%5B0%5D=RIO&dateFrom%5B0%5D=17-03-2017&dateTo=31-03-2017&adults=1&children=0&babies=0&flightClass=NMO.GBL.SCL.ECO&hostname=www.avantrip.com

- Factor 18: abstinencia de enlaces rotos. Este factor hace referencia a evitar que permanezca la existencia de URL que ya no existan o que ya no tengan contenido. Una de las situaciones más perjudiciales para un sitio web, a nivel orgánico, es que el enlace indexado en el buscador sea un link roto que lleva a una página de error. Si, por alguna razón, el contenido de la página en cuestión fue dado de baja, es importante redireccionar al usuario a una página equivalente que sí exista.
- Factor 19: utilización de parámetro *no index-no follow*. Es muy común que, en algunos momentos en los que una página está siendo modificada o se está preparando un nuevo sitio web, a este nuevo contenido se le coloquen etiquetas *no index* y *no follow* para evitar su indexación cuando todavía no está listo. Esto es correcto, pero, en muchas oportunidades, esta etiqueta es olvidada y queda puesta en el sitio web al momento de su publicación, lo que causa, obviamente, un problema de contenido no indexado por los robots de búsqueda. Esto es fácilmente detectable y solucionable al retirar dichas etiquetas.
- Factor 20: abstinencia de nombres genéricos en las URL. Este factor está muy ligado a las URL amigables, ya que, si bien aquí no se está hablando de parámetros ilegibles para la lectura natural, sí sucede que en varias oportunidades existen nombres tan generales que no aportan nada orgánicamente. Si bien esto no resta puntos de relevancia, sí está dejando de ganar lo que podría obtener si se utilizaran palabras clave apropiadas en la estructura de la URL. Aquí, un ejemplo:

- **Sitio web con URL genéricas:**

www.dominio.com/paquete/opcion-1

www.dominio.com/paquete/opcion-2

www.dominio.com/paquete/opcion-3

- **El mismo sitio web si tuviera URL con nombres específicos:**

www.dominio.com/paquetes-de-viaje/rio-de-janeiro-brasil

www.dominio.com/paquetes-de-viaje/amsterdam-holanda

www.dominio.com/paquetes-de-viaje/sydney-australia

- Factor 21: manejo de ID de usuario o sesión. En la medida de lo posible, se recomienda que los números de identificación de la sesión o el usuario que está navegando el sitio se coloquen al final de la URL como un parámetro posterior a un `?` o a un `#`, ya que estos parámetros pueden ser utilizados para indicar que la URL y su contenido llegan hasta allí y posterior a esos parámetros que el buscador no los tome en cuenta. Esta solución es similar a la forma de manejar las URL dinámicas, ya que permite remediar los requerimientos a las bases de datos, al mismo tiempo que se puede declarar al *bot* del buscador que no se continúe leyendo a partir de ese parámetro y reconoce el contenido como parte de una sola página.
- Factor 22: abstinencia de repetición de palabras clave en la URL. Como dice el dicho: *menos es más*. Por eso, si ya se utilizó una palabra clave en la URL, se recomienda no duplicarla. Ya de por sí el espacio disponible para optimizar los largos de las URL son cortos. No vale la pena repetir palabras que ya están colocadas y, más bien, se podría sacar mayor potencial al utilizar esos caracteres para otro tipo de *keyword* (palabra clave) que pudiera ampliar el mercado al que se enfoca, que podría encontrar dicha página.
- Factor 23: abstinencia de duplicidad de accesos por URL distinta. Esto hace referencia a evitar que distintas URL lleven al mismo contenido debido a que, de esta manera, podrían ser considerados contenidos duplicados o incluso plagio. Este es uno de los factores más penalizados por parte de los algoritmos de Google.
- Factor 24: abstinencia de utilización de mayúsculas en la URL. En ningún caso deben ser utilizadas mayúsculas en la URL. Sin excepción.
- Factor 25: separación de palabras en la URL. Para separar las palabras en la URL, se recomienda utilizar guion medio. De esta manera, se mantendrá una lectura cómoda, al mismo tiempo que se evitarán caracteres incómodos que ensuciarían la URL. Por ejemplo, para un contenido llamado “Paquetes a Río de Janeiro”, la URL podría quedar de la siguiente manera:

www.dominio.com/paquetes-rio-de-janeiro

Si se intentara dejar los espacios, eventualmente la URL sería interpretada así:

www.dominio.com/paquetes%20rio%20de%20janeiro

Los espacios serían reconocidos automáticamente con el patrón “%20”, lo que ensucia la URL.

1.1.4 Importancia del mapa de sitio, robots y similares

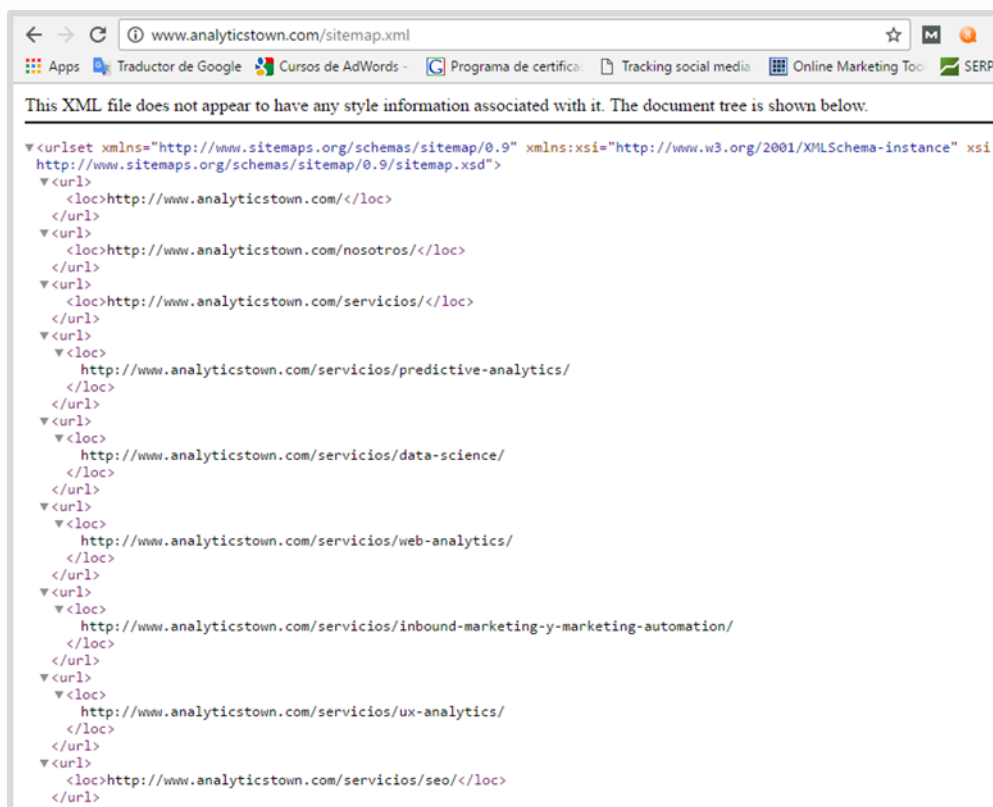
Hay elementos y factores que afectan la optimización de una estrategia SEO desde instancias simples pero importantes a las que no siempre se les da la atención que merecen. Esta sección está destinada a esos factores.

- Factor 26: manejo de mapa de sitio en HTML. Es una buena práctica mantener una sección en su sitio web que funcione como un mapa de sitio en HTML para que los usuarios y los robots de búsqueda puedan acceder a él cuando necesiten ubicarse donde están o quieran acceder fácilmente a otra sección que se encuentra en otro nivel de contenido. En los últimos años, también se utiliza mucho la modalidad de colocar un mapa de sitio de las secciones principales al pie de cada página.
- Factor 27: definición de dominio preferido. Este factor está dedicado a definir a través de un redireccionamiento, así como a través de una declaración en Google Search Console, de qué forma se prefiere que se utilice el dominio. En la gran mayoría de casos, la disyuntiva pasa por si se utiliza con `www` o sin ello. Por lo general, para cumplir con las normas de los buscadores, en este caso se pueden aplicar dos acciones en paralelo para que esto siga su curso: la primera es un redireccionamiento 301 de la versión sin `www` a la que sí lo tiene y la segunda es al declarar en Google Search Console que el dominio preferido es con `www`.
- Factor 28: uso de sitemap XML. Un `sitemap.xml` es un archivo que deberá estar ubicado en la siguiente dirección por defecto: `www.dominio.com/sitemap.xml`

En este archivo, se declara la totalidad de las URL del sitio web y se puede, incluso, declarar la prioridad de las páginas en una escala de 0 a 1 al colocar los intermedios en decimales según se le desee asignar mayor o menor prioridad a cada contenido. También es posible declarar la periodicidad con la que se desea que el navegador del buscador pase al revisar cada página del sitio web según como se considere que se tendrá actualizaciones para mostrar. Tanto la prioridad como el tiempo para que el robot del buscador pase revisando el sitio son parámetros opcionales que pueden omitirse sin problema. Cuando se omiten, el buscador será quien decida qué prioridad asignar a cada página en cada búsqueda que las active y cada cuánto tiempo pasará visitando las páginas, según la periodicidad con la que vaya notando que tengan contenido fresco.

El siguiente es un ejemplo de cómo se ve un archivo `sitemap.xml` en su respectiva ubicación en un sitio web:

Figura 5. Ejemplo de Sitemap XML



```
<?xml version="1.0" encoding="UTF-8" ?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:
http://www.sitemaps.org/schemas/sitemap/0.9/sitemap.xsd">
  <url>
    <loc>http://www.analyticstown.com/</loc>
  </url>
  <url>
    <loc>http://www.analyticstown.com/nosotros/</loc>
  </url>
  <url>
    <loc>http://www.analyticstown.com/servicios/</loc>
  </url>
  <url>
    <loc>http://www.analyticstown.com/servicios/predictive-analytics/
    </loc>
  </url>
  <url>
    <loc>http://www.analyticstown.com/servicios/data-science/
    </loc>
  </url>
  <url>
    <loc>http://www.analyticstown.com/servicios/web-analytics/
    </loc>
  </url>
  <url>
    <loc>http://www.analyticstown.com/servicios/inbound-marketing-y-marketing-automation/
    </loc>
  </url>
  <url>
    <loc>http://www.analyticstown.com/servicios/ux-analytics/
    </loc>
  </url>
  <url>
    <loc>http://www.analyticstown.com/servicios/seo/</loc>
  </url>
</urlset>
```

Fuente: Captura de pantalla de Analytics Town, 2016, <https://bit.ly/3bgUW3s>

- Factor 29: carga de sitemap XML a Google Search Console. El mismo archivo sitemap.xml que se publicó en su sitio web debe ser cargado a la herramienta Google Webmaster Tools para notificarle a Google que un nuevo mapa de sitio ha sido actualizado y que sepa qué URL debe comenzar a revisar para indexarlas en su listado.
- Factor 30: carga de sitemap XML a Bing Webmaster Tools. De la misma forma que se hace la carga del *sitemap* a Google Search Console, también se puede suministrar a la herramienta Bing Webmaster Tools. Con ello, se hará el mismo proceso de notificación al buscador de Bing para que este sea reconocido e indexado por este buscador.
- Factor 31: uso de Robots.txt. De la misma forma que el archivo sitemap XML se publica en una dirección específica por defecto en el sitio, también el archivo Robots.txt tiene su propia ubicación y funcionalidad.

Un archivo robots.txt es una especie de manual de instrucciones para los robots del buscador en el que se declara qué contenido es indexable y cuál no para cada tipo de *bot*, ya que existen diferentes nombres de robot a los que se les puede brindar indicaciones de qué incluir o excluir. Este archivo deberá ser un archivo de texto ubicado exactamente en la siguiente URL: www.dominio.com/robots.txt

- Factor 32: uso de funciones *allow* y *disallow* en Robots.txt:
La instrucción básica por defecto para un archivo robot.txt es la siguiente:

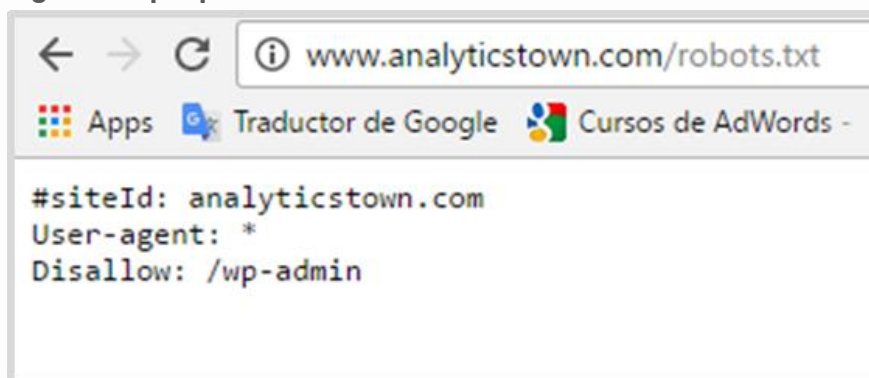
User-agent: *

Disallow: {*parámetros a excluir*}

En ella se está indicando que cualquier *user agent* (como podrían ser el robot de Google para *desktop*, el robot de Google para *mobile*, el de Bing para *desktop* y *mobile*, los de Yahoo, etcétera) reconozcan como ubicaciones de URI a no incluir o deshabilitar con la instrucción *disallow*.

A continuación, se presentará un ejemplo de cómo podrían verse dos tipos de robots.txt, uno simple básico y otro con gran cantidad de elementos a excluir de la indexación en los buscadores.

Figura 6: Ejemplo de robot.txt común



```
#siteId: analyticstown.com
User-agent: *
Disallow: /wp-admin
```

Fuente: Captura de pantalla de *Analytics Town*, 2016, <https://bit.ly/3bgUW3s>

Figura 7. Ejemplo de robot.txt con varios parámetros de exclusión

```
User-agent: *
Crawl-delay: 10
# Directories
Disallow: /includes/
Disallow: /misc/
Disallow: /modules/
Disallow: /profiles/
Disallow: /scripts/
Disallow: /themes/
# Files
Disallow: /cron.php
Disallow: /update.php
Disallow: /xmlrpc.php
Disallow: /boost_stats.php
# Paths (clean URLs)
Disallow: /admin/
Disallow: /comment/reply/
Disallow: /logout/
Disallow: /node/add/
Disallow: /search/
Disallow: /user/register/
Disallow: /user/password/
Disallow: /user/login/
# Disallow: /beneficios/
Disallow: /consultas-y-sugerencias/
Disallow: /accessbanking/
Disallow: /beneficios-filtro
Disallow: /beneficios-filtro/

# Paths (no clean URLs)
Disallow: /?q=admin/
Disallow: /?q=comment/reply/
Disallow: /?q=filter/tips/
Disallow: /?q=logout/
Disallow: /?q=node/add/
Disallow: /?q=search/
Disallow: /?q=user/password/
Disallow: /?q=user/register/
Disallow: /?q=user/login/
```

Fuente: Elaboración propia.

- Factor 33: uso de metarrobots. Los metarrobots son etiquetas que se colocan en el HTML de una página en particular que se desee bloquear a la indexación de los buscadores a través de una meta etiqueta en el *body*.

Esta metaetiqueta tiene un efecto similar al que tendría la instrucción *disallow* del archivo robots.txt con la diferencia que va colocada en el código de la página que se desea excluir.

La metaetiqueta robot tiene la siguiente sintaxis:

```
<meta name="robots" content="noindex, nofollow">
```

En este ejemplo se está indicando que la página en donde esta instrucción fue incluida no debería ser indexada ni tampoco seguida.

Esto aplica para cuando se tiene un contenido que se desea dejar privado o, por ejemplo, para un nuevo sitio previo a su lanzamiento oficial para evitar duplicación de contenido mientras aún no se desea que se publique. En este último caso, no hay que olvidar retirar dichas etiquetas cuando haga el lanzamiento. De lo contrario, esta instrucción seguirá vigente y verá que el sitio no se indexará en los buscadores.

- Factor 34: uso de etiquetas canónicas. Como parte de los elementos que brindan calidad al contenido, se encuentran las etiquetas *canonicals* dado que, a través de ellas, es posible declarar ante los motores de búsqueda que un contenido posee distintas versiones, ya sea dentro de un mismo sitio web o republicado en múltiples sitios.

Con la etiqueta *canonicals*, se define cuál de todas las versiones es la original y se establece a quién debe darse el crédito por dicho contenido, inicialmente. Se recomienda agregar, dentro del body de cada html de contenido una etiqueta con la siguiente estructura:

Para que esta estructura sea correcta, se plantea el siguiente ejemplo:

```
<link rel="canonical" href="{URL de contenido original}">
```

Si un mismo contenido aparece en las páginas:

www.dominio.com/pagina1 (original)

www.dominio.com/pagina2

En las versiones anteriores de URL, asumiendo que el contenido sea el mismo, pero esté reproducido o republicado por necesidades editoriales, se debe agregar una etiqueta canónica como la siguiente:

```
<link rel="canonical" href="https://www.dominio.com/pagina1">
```

Se recomienda agregar esta etiqueta, referenciando a todo el contenido que sea original y no duplicado. Esta etiqueta también se colocará en la página original, autorreferenciándose.

Conceptos clave

- Meta títulos, metadescripciones y URL son los tres factores que más visualización reciben por ser parte de los elementos que componen el *snippet* del resultado de búsqueda.
- Existen normas y estándares que deben respetarse para cumplir con las buenas prácticas de SEO y, según ello, podrá optimizarse el sitio web deseado.

- Es importante tomar en cuenta tanto los factores básicos visibles como la URL, títulos y descripciones, así como los factores más técnicos que requieren un tratamiento especial como el *sitemap*, los robots y otros tipos de elementos más técnicos.

Referencias

Analytics Town. (2016). [Captura de pantalla]. Recuperado de <http://www.analyticstown.com/sitemap.xml>

Analytics Town. (2016). [Captura de pantalla]. Recuperado de <http://www.analyticstown.com>