



چند نکته مهم در رابطه با سئو (جوملا)

نویسنده: فرهاد رضا زاده

www.modironline.com

انتشار یافته و تأیید شده توسط

Joomina.ir

مرکز داندلود افزونه های تخصصی جوملا

مقدمه نویسنده

چیزی که به نظر می رسد این است که بهینه سازی سایت برای موتورهای جستجو عملی پیچیده و با هزینه بالا می باشد ولی باید خدمت شما عرض کنم که این قضیه برای سایت های جوملایی کمی متفاوت است. علت این امر هم قدرت بالای سیستم مدیریت محتوای جوملا می باشد که به شما این امکان را می دهد تا تعداد زیادی صفحه با کیفیت بالا برای سایت خود ایجاد نمایید. بنابراین شناخت و تسلط بر سیستم مدیریت محتوای جوملا این امکان را به شما می دهد راحت تر از چیزی که به نظر شما می رسد سایت خود را به صورت **مقدماتی** برای موتورهای جستجو بهینه سازی نمایید. در این مقاله سعی شده بدون اشاره به افزونه های فوق العاده قدرتمندی که برای سئو جوملا وجود دارد (و شخصا به آن ها اعتقادی ندارم) برخی از نکاتی را که تا بحال کم تر کسی به آن اشاره کرده را ذکر نمایم. نکاتی به نظر ساده ولی فوق العاده پر اهمیت. در این مقاله کوتاه سعی نموده ام هم افرادی که با سئو آشنا نیستند و هم کسانی که در این زمینه مهارت دارند را در نظر بگیرم. در کار حرفه ای ما افرادی که به صورت حرفه ای روی سئو کار می کنند و اطلاعاتی را در این زمینه از منابع معتبر کسب می کنند تا زمانی که اطلاعات خود را به صورت عملی آزمایش نکرده و نتایج آن عمل را کسب نکرده اند نباید ادعای این موضوع را داشته باشند که فلان عمل برای سئو سایت مناسب است (به جز چیزهای بدیهی و گفته شده توسط خود موتورهای جستجو) و برای سایت های جوملایی که جای خود دارد. و همانطور که می دانید برای تست کردن عملی مربوط به سئو سایت شاید تا چندین ماه زمان نیاز باشد. همین موضوع است که تجربه در سئو را نسبت به علم آن برتری می دهد. در رابطه با اینکه چرا تا به حال انجمن قدرتمندی در رابطه با سئو در ایران نداریم دلایل متعددی وجود دارد. اول اینکه متخصصان این امر که هم دارای

تجربه و هم دارای دانش باشند در ایران خیلی کم می باشند و دوم اینکه چنین انجمنی نیاز به یک کار تیمی دارد و از عهده ی یک نفر و دو نفر خارج می باشد سند حرف من هم انجمن هایی است که در حال حاضر دارند فعالیت انجام می دهند ولی آنطور که باید موفق باشند، موفق نیستند. انشاالله که جامعه وب کشور از این اوضاع بیرون بیاید و به مرحله ای برسیم که هیچ موتور جستجویی اجازه تحریم سایت های ایرانی را به خود ندهد. (در صورتی که دوستان طراحان وب ابراز تمایل نمایند مقاله ای در زمینه رابطه سئو با زبان برنامه نویسی PHP که سخت و پیچیده هم می باشد، نوشته و در اختیارشان قرار می دهم. در صورت هر گونه سوال می توانید از طریق سایتیم با من ارتباط برقرار نمایید: www.modironline.com) امیدوارم این **نکات کوتاه** برای شما که کار با جوملا را به صورت حرفه ای دنبال می کنید مفید باشد.

در انتخاب کلمات کلیدی حتما به نکات زیر توجه داشته باشید :

- قبل از انتخاب کلمات کلیدی چند دقیقه فکر کنید و خود را به جای کاربر هدف قرار دهید به گونه ای که اگر شما به جای او باشید و قرار باشد در موتورهای جستجو واژه ای جستجو کنید چه واژه ای جستجو می کنید.
- کلمات کلیدی شما به گونه ای انتخاب شوند که شامل نام برند، محصولات، خدمات و اطلاعات شما باشند.
- با توجه به اینکه اکثر موتورهای جستجو حساس به حروف بزرگ نیستند نیازی به استفاده از حروف بزرگ در واژه های خود نیستید.
- نباید در کلمات کلیدی از کلماتی استفاده نمایید که دارای املای یکسان ولی معانی متفاوت هستند زیرا ممکن است کاربر با یکی از این واژه ها وارد سایت شما شده ولی به اطلاعات مورد نیاز خود دسترسی پیدا نکند. موتورهای جستجو بعضی اوقات نتایجی را به شما نمایش می دهند که مربوط به واژه ی جستجو شده توسط شما می باشد که استفاده از این واژه های مرتبط می تواند به عنوان کلمه کلیدی مناسب شناخته شود.

هنگام نصب جوملا نباید اطلاعات نمونه آن را بر روی سایت خود نصب نمایید:

یکی از مهم ترین دلایلی که توصیه می شود که اکیدا اطلاعات نمونه را نصب نکنید این است که اگر این اطلاعات در پایگاه داده موتورهای جستجو مانند گوگل ایندکس شوند از دو جنبه سئو سایت شما با مشکل روبه رو می شود و هم در رتبه و هم پیج رنک شما تاثیر می گذارد. اول اینکه موتورهای جستجو اطلاعاتی از سایت شما را ثبت نموده اند که هیچ ارتباطی با ماهیت و موضوع سایت شما ندارد و دلیل دوم که مهمتر نیز می باشد این است که پس از حذف این اطلاعات نمونه از سایت خود آدرس ها و لینک هایی را که در موتورهای جستجو ثبت شده اند از بین برده اید که این امر خود بر روی سئو سایت به ویژه میزان اعتمادی که موتورهای جستجو باید نسبت به سایت شما پیدا کنند تاثیر به شدت منفی دارد. پس توصیه می شود که از نصب اطلاعات نمونه خودداری نمایید.

غیر فعال کردن لینک مربوط به pdf ها



زمانی که که ما سایت جوملایی را بررسی می کنیم گاهی اوقات متوجه می شویم که لینک های حاوی فایل های pdf در رتبه بالاتری نسبت به صفحات اصلی شما در موتورهای جستجو قرار دارند. حال اشکال این اتفاق کجاست؟ کاربران قادر نیستند که به صورت مستقیم از طریق فایل های pdf وارد سایت شما شده و از تمام اطلاعات و خدمات سایت شما استفاده نمایند. پس شما می توانید در قسم مدیریت مطالب خود لینک مربوط به pdf ها را غیر فعال نمایید.

قفل کردن آدرس سایت بر روی www

قبلا در برخی از انجمن ها مطلبی را در این زمینه مطالعه کردم که جا دارد از تهیه کنندگان آن تشکر کنم. من در این قسمت سعی می کنم آن مطلب را کامل نمایم. بعضی ها فکر می کنند که داشتن و نداشتن www موجب ایجاد دو آدرس برای موتورهای جستجو می شود این در حالی است که اگر یک سایت جوملایی داشته باشید و از پروتکل https نیز صرف نظر کنیم حداقل دارای 4 آدرس به شرح زیر است:

<http://yoursite.com>

<http://yoursite.com/index.php>

<http://www.yoursite.com>

<http://www.yoursite.com/index.php>

شکل معمول و استاندارد آن می شود از آن برای معرفی آدرس صفحه اصلی سایت خود استفاده نماییم به صورت:

<http://www.yoursite.com>

می باشد. اهمیت این موضوع در این است که اگر صفحه ی اول شما چهار آدرس متفاوت داشته باشد خیلی از پارامترهای مهم مثل میزان بازدید سایت شما و همچنین لینک هایی که به سایت شما داده می شود روی این چهار آدرس تقسیم می شود که اصلا برای سئو شما مناسب نیست. و برای این منظور بایستی قطعه کد زیر را به فایل htaccess خود اضافه نمایید:

RewriteEngine On

index.php redirect

RewriteCond %{THE_REQUEST} ^[A-Z]{3,9}\ /index\.php HTTP/

RewriteRule ^index\.php\$ <http://www.yoursite.com/> [R=301,L]

Redirect non-www to www version

RewriteCond %{HTTP_HOST} ^yoursite\.com\$ [NC]

RewriteRule ^(.*)\$ <http://www.yoursite.com/>\$1 [R=301,L]



البته این موضوع به این سادگی هم که به نظر می رسد نیست. اگر شما سایت جوملایی داشته باشید از طریق آدرس زیر نیز می توانید به صفحه اصلی خود دسترسی داشته باشید:

www.yoursite.com/component/content/frontpage

اگر با این آدرس به سایت خودتان وارد شوید به فاجعه بودن آن برای کاربران پی خواهید برد.

جالبه بدانید که این آدرس جز اولین آدرس هایی است که از سایت شما در موتورهای جستجو ایندکس می شود. برای حل این مشکل و مشکلات شبیه به این موضوع می توانید از دستور زیر استفاده نمایید:

Redirect 301 / component/content/frontpage <http://www.yoursite.com>

با توجه به اینکه نیازی به توضیح دستور بالا ندارید از توضیح آن صرف نظر می کنم.

فعال کردن cache :

سایت هایی با حجم کم و سرعت بالا هم برای کاربران و هم برای موتورهای جستجویی مانند گوگل فوق العاده حائز اهمیت می باشد و این امکان را هم به کاربران و هم به موتورهای جستجو می دهد که بهتر و مفیدتر با سایت شما ارتباط برقرار نمایند. برای این منظور می توانید در قسمت تنظیمات عمومی سایت خود **cache** را فعال نمود ولی قابل خواندن و نوشتن بودن فایل **cache** خود را نیز به یاد داشته باشید.

Spiderها و robot ها و crawler ها و سایت های جوملایی

شاید شما قبلا اطلاعاتی در زمینه Spiderها و robot ها و crawler ها داشته باشید و بدانید که آن ها به گونه ای برنامه نویسی می شوند که عمل **crawl** کردن محتویات سایت شما را مثل متن ها و لینک ها را انجام دهند و آن ها را در دیتابیس موتورهای جستجو قرار داده و با توجه به الگوریتم موتورهای جستجو سایت شما را رتبه بندی نمایند. جالبه بدانید که **crawler**ها وقتی برای اولین بار وارد سایت شما می شوند به اولین چیزی که توجه می کنند لینک های شما در صفحه هایتان می باشد. سپس متن های شما را می خوانند و لینک هایی که قبلا بررسی نموده اند دریافت می کنند.

این که دقیقا گفته شود که یک **crawler** چگونه اطلاعات یک سایت را مرور می کند کمی دشوار تر از اینه که من توضیح دادم. **Crawler** یک درخواست ارسال می کند به یک سروری که از سایت شما میزبانی می کند و نحوه دریافت این درخواست تقریبا مثل حالتی که مرورگرها صفحاتی را که شما می خواهید ببینید پیگیری می کنند.

Crawler می توانند در هر زمانی صفحات سایت شما را مورد بررسی قرار دهند. همین موضوع به ویژه در سایت های پربازدید ایجاد مشکل می کند که مهمترین آن این است که سرعت لود سایت را پایین می آورد و موجب از دست دادن



ترافیک شما در اوج ترافیک سایتتان می شود. و حتی ممکن است در ساعاتی که سایت شما دارای بازدید زیادی است Crawler سایت شما را ترک کند. هنگام برنامه نویسی و طراحی یک سایت باید به این نکته توجه داشت که دیر لود شدن سایت امکان جریمه شدن سایت ما توسط Crawler ها به وجود می آورد ، اما برای این موضوع راه حلی هم وجود دارد و آن این است که کاری انجام دهیم که Crawler در سایت خود نگهداریم. من به طور خلاصه به شما خواهیم گفت چه کاری انجام دهید که Crawler مجدداً به سایت شما راهنمایی شوند ولی اگر تجربه من را می خواهید در صورتی که لزومی ندارد که سایتی سنگین جهت لود شدن داشته باشید بهتر است این کار را انجام ندهید و اگر غیر از این باشد راه حل مناسب را انتخاب نمایید. ولی قبل از اینکه وارد این بحث شوم به شما یک سری از مواردی را خواهیم گفت که نباید اجازه دهید توسط Crawler ایندکس شوند و حتی از آنها در طراحی سایت خود استفاده نمایید:

زمانی که در حال ایجاد یک صفحه هستید:

هنگامی که در حال ایجاد و ویرایش یک صفحه هستید مطمئن شوید که آن صفحه جزو صفحاتی نیست که توسط Crawler مورد بررسی می باشد و از این قضیه نیز نگران نباشید زیرا پس از آماده شدن صفحات شما توسط Crawler مورد بررسی قرار خواهد گرفت.

صفحات حاوی لینک های زیاد:

این گونه از صفحات که حاوی لینک های زیادی هستند (به استثنا نقشه سایت) برای Crawler ها مشکوک محسوب می شوند و حتی لازم است که این نکته را خدمت شما عرض کنم که ما این گونه از سایت ها را در زمره سایت های سپم قرار می دهیم حتی اگر موتورهای جستجو این موضوع را تشخیص ندهند. برای حل این مشکل می توان از توضیحات و متن به جای بکاربردن لینک های زیاد به تنهایی استفاده نماییم.

صفحاتی که دارای محتویات خیلی قدیمی هستند:

محتویات قدیمی سایت شما مانند آرشیو وبلاگ شما ضرورتاً آسیبی به رنک شما در موتورهای جستجو وارد نمی کند ولی نبود آن ها موجب ارتقای رنک شما خواهد شد. موضوع آزاردهنده ای که در رابطه با آرشیو ها با آن رو به رو هستیم تعداد دفعاتی است که این محتویات بر روی سایت شما آشکار و لود می شوند. و همچنین ممکن است این اشتباه برای Crawler ها پیش آید که شما از محتویات یکسان چندین بار در سایت خود استفاده نموده اید که این خود مصداق سپم برای سئو می باشد. پس بهتره که ریسک نکرده و آرشیو خود را محدود کرده و مطمئن شویم که محتویاتی که گذشته به کار برده اید مجدداً به کار نبرید.

اطلاعات شخصی:



بهرتر این است که اطلاعات شخصی را بر روی وب منتشر نکنیم ولی اگر به هر دلیلی مجبور هستید این کار را انجام دهید امکان دسترسی Crawler را به آن ها ببندید. دلایل زیادی وجود دارد که نباید اجازه دهید در بعضی از مواقع Crawler ها صفحات شما را ایندکس نمایند و مواردی که من عرض نمودم تعدادی از رایج ترین این موارد بود.

فایل robots.txt

Crawler ها این توانایی را دارند که به یک وب سایت آسیب وارد نمایند. ولی با انجام یک سری اقدامات می توان رفتار آن ها را کنترل نمود. این اقدامات به وسیله robots.txt صورت می گیرد. فایل robots.txt یک عنصر واقعی ، مهم و تاثیر گذار است که شما بر روی آن کنترل دارید. این فایل در root سایت شما قرار می گیرد. این فایل حاوی دستور العمل های ضروری و مهمی می باشد که وظیفه کنترل Crawler ها را بر عهده دارد و به آن ها می گوید چه قسمتی از سایت شما را ایندکس کنند و چه قسمت هایی را ایندکس نکنند. برای ایجاد ارتباط با Crawler ها باید از دستورالعملی خاصی استفاده نمایید که آن ها متوجه شوند. در بیش تر فرم های این فایل از دستور زیر استفاده می کنیم:

User-agent: *

Disallow: /

این دو قسمت خیلی ضروری و مهم می باشند. قسمت اول یعنی User-agent: بیانگر فرمان شما به Crawler ها می باشد. نشان ستاره (*) نشان می دهد که تمام و انواع Crawler ها را پوشش می دهد ولی این قابل کنترل و تغییر توسط شما می باشد و می توانید تنها به یک یا چند Crawler اجازه ارتباط دهید.

قسمت دوم یعنی Disallow: به Crawler می گوید که به چه قسمت هایی اجازه دسترسی ندارد. و علامت اسلش (/) نشان دهنده تمام دایرکتوری ها می باشد. پس این قسمت به Crawler می گوید از ایندکس کردن دایرکتوری ها چشم پوشی کن.

علامت (:) بیانگر این موضوع است که Crawler ها با دستور مهمی روبه رو هستند و باید به آن توجه کنند. معمولا از Crawler ها خواسته می شود که از اکثر دایرکتوری ها صرفه نظر کنند. به عنوان مثال اگر بخواهید که به Crawler ها اعلام نمایید که از دایرکتوری های موقتی چشم پوشی نماید به صورت زیر عمل می نمایم:

User-agent: *

Disallow: /tmp/

و شاید شما بخواهید از دایرکتوری های بیشتری چشم پوشی شود:

User-agent: *

Disallow: /tmp/

Disallow: /private/



Disallow: /links/listing.html

در دستور بالا از Crawler ها خواسته می شود که از دایرکتوری های موقت و شخصی و همچنین صفحاتی که شامل لینک ها است چشم پوشی نمایند. نکته ای که همیشه باید در ذهن داشته باشید این است که Crawler ها فایل robots.txt را از بالا به پایین می خوانند پس اگر شما می خواهید یک سری فرمان بنویسید باید به چگونه نوشتن آن دقت کنید.

به مثال زیر توجه کنید:

User-agent: *

Disallow: /tmp/

User-agent: CrawlerName

Disallow: /tmp/

Disallow: /links/listing.html

اگر توجه کنید می بینید که در یک فرمان به تمام Crawler ها اجازه دسترسی و ارتباط داده اید و در فرمانی دیگر تنها به یک Crawler خاص. در این جا آن Crawler خاص فرمان را نخواهد انجام داد زیرا قبل از آن شما اجازه ارتباط به تمام Crawler ها را داده اید.

صحیح دستور بالا به صورت زیر می باشد:

User-agent: CrawlerName

Disallow: /tmp/

Disallow: /links/listing.html

User-agent: *

Disallow: /tmp/

User-agent: CrawlerName

Disallow: /tmp/

Disallow: /links/listing.html

User-agent: *

Crawler های موتورهای جستجو متفاوت دارای نام های متفاوتی هستند که اگر شما log وب سرور را بررسی کنید احتمالا نام آن ها را مشاهده خواهید نمود. برخی از آن ها در لیست زیر آورده شده است:

Google: Googlebot

MSN: MSNbot

Yahoo! Web Search: Yahoo SLURP or just SLURP

Ask: Teoma

AltaVista: Scooter

LookSmart: MantraAgent



WebCrawler: WebCrawler
SearchHippo: Fluffy the Spider

لیست بالا تعداد اندکی از Crawler های هستند که باید سایت شما را مورد بررسی قرار دهند. حال شاید بپرسید دانستن اینکه هر کدام از Crawler ها مربوط به چه موتور جستجویی هستند چه فایده ای دارد من در جواب باید به شما عرض کنم که ما یکسری spambots و Crawler مخرب داریم که معمولا به یکسری دلایل معمولا بی اخلاقانه دوست دارند با سایت ها به ویژه سایت های مهم و حساس ارتباط برقرار نمایند. اگر شما نام آن ها را بدانید به آن ها اجازه ارتباط با سایت خود و همچنین بررسی اطلاعات شخصی و حساس خود را نخواهید داد. البته تضمین صددرصدی هم وجود ندارد.

متا تگ Robots:

وقتی که از متا تگ Robots استفاده می نمایید این امکان را برای Crawler ها فراهم می نمایید که به جای استفاده از فایل robots.txt از این تگ استفاده نمایند و با سایت شما ارتباط برقرار نمایند. متا تگ robots قطعه کدی به صورت HTML می باشد که در تگ <HEAD> سایت شما قرار می گیرد. این متا تگ برای کسانی که می خواهند از داخل سایت Crawler ها را کنترل نمایند مناسب می باشد.

به مثال زیر توجه نمایید:

```
<html>
<head>
<meta name="robots" content="noindex, nofollow">
<meta name="description" content="page description.">
<title>
Web Site Title
</title>
</head>
<body>
```

کد بالا بیانگر این می باشد که Crawler ها نباید محتویات سایت را ایندکس نمایند و نباید لینک های سایت را دنبال نمایند. به مثال های زیر نیز توجه نمایید:

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

تفاوت عمده ای که بین فایل robots.txt و متا تگ های robots وجود دارد این است که متا تگ robots این امکان را برای ما فراهم نمی کند که Crawler خاصی را انتخاب نماییم و همچنین در برابر با robots.txt دارای دقت خیلی کم تری می باشد.

خیلی از Crawler ها هستند که نه فایل robots.txt و نه متا تگ های robots را می شناسند و شما نیز نمی توانید کنترلی بر روی آن ها داشته باشید. همانطور که Crawler ها می توانند موجب ایندکس شدن صفحات شما شوند ، همچنین می توانند در صورتی که درست کنترل و هدایت نشوند موجب آسیب به سایت شما شوند.



فایل استاندارد robots.txt برای جوملا :

به صورت استاندارد فایل robots.txt برای جوملا به صورت زیر می باشد:

```
User-agent: *  
Disallow: /administrator/  
Disallow: /cache/  
Disallow: /components/  
Disallow: /images/  
Disallow: /includes/  
Disallow: /installation/  
Disallow: /language/  
Disallow: /libraries/  
Disallow: /media/  
Disallow: /modules/  
Disallow: /plugins/  
Disallow: /templates/  
Disallow: /tmp/  
Disallow: /xmlrpc/
```

همانطور که مشاهده می کنید بعضی از دایرکتوری های خاص برای موتورهای جستجو بلوکه شده اند.

یک نکته خیلی مهم در رابطه با دستور Disallow: /images/ :

به صورت استاندارد دایرکتوری مربوط به تصاویر برای موتورهای جستجو بلوکه شده است. با این کار شما این امکان را به موتورهای جستجو نخواهید داد تا تصاویر مربوط به سایت شما را ایندکس نمایند. این نکته خیلی مهمی می باشد که اکثر سایت های جوملایی در ایران آن را رعایت نمی کنند. یعنی اگر شما لوگو یا برند تجاری داشته باشید و در سایت خود استفاده کرده باشید در موتورهای جستجو ثبت نمی شود. البته زمانی شما می توانید این کار را انجام دهید که تصاویر شما بهینه سازی شده باشند. بدین گونه که باید تصاویر شما دارای نام باشند. همچنین می بایست از تگ alt برای زبان فارسی استفاده نمایید. این کار چند حسن دارد، مثلا اگر کاربر نتواند تصاویر را در سایت شما ببیند می تواند متن داخل alt را مشاهده کرده و مثلا این اطمینان برایش حاصل شود و بر روی تصویر کلیک نماید. مزیت دیگری هم که دارد این است که اگر تصویر به جایی لینک داده شده باشد متن داخل alt یک جورایی حکم anchor text دارد. البته خود متنی که در تگ alt قرار می گیرد باید شرایطی خاص داشته باشد (از دست استاد تهرانی) مثلا این متن طولانی نباشد و از کلمات کلیدی سایت به صورت مکرر در آن ها استفاده نشود. ولی توصیه می شود از تعداد زیادی از تصاویر برای این منظور استفاده نشود. اگر تصاویر شما به این گونه باشند پس بهتر است که دستور Disallow: /images/ از فایل robots.txt خود حذف نمایید. اگر شما SEF patch را بر روی سایت خود نصب نمایید این کار را برای شما انجام می دهد.

فایل xml نقشه سایت :

فایل xml نقشه سایت در واقع همراه و همکار فایل robots.txt محسوب می شود. با استفاده از این فایل شما به Crawler ها خواهید گفت که چه قسمت هایی در سایت شما برای ایندکس کردن آن ها وجود دارد. در حال معمول این فایل

شامل تمام آدرس های موجود در یک وبسایت می باشد. دلیل استفاده از این فایل این است که صفحات واقعی سایت خود را به Crawlerها معرفی نموده و اطمینان لازم جهت ایندکس شدن این مطالب برای آن ها کسب شود. این فایل در همان مکانی که فایل robots.txt قرار گرفته قرار می گیرد. و شما می توانید آن را در وبستر یا دایرکتوری های موتورهای جستجو ثبت نمایید. از دیگر مزیت های دیگر استفاده از نقشه سایت این است که موجب می شود که موتورهای جستجو صفحات شما را زودتر در دیتابیس خود ایندکس نمایند. البته این موضوع را هم به شما عرض کنم که استفاده از این فایل تضمینی برای ایندکس شدن تمام صفحات شما را نمی دهند و تنها به عنوان یک راهنما برای Crawlerها عمل می نماید. سایت های زیادی هستند که به صورت رایگان این امکان را به شما می دهند تا بتوانید نقشه سایت خود را ایجاد نمایید ولی به این دلیل شاید دوستانی بخواهند به صورت حرفه ای از کم و کیف آن ها مطلع شوند من به آموزش آن می پردازم:

چگونه شخصا فایل xml نقشه سایت خود را ایجاد نمایید :

نقشه سایت شما احتیاج به دو بخش اصلی دارد.

1- نقشه شما باید شامل تگ های XML باشد

2- نقشه شما باید بر اساس UTF-8 نوشته شود و نوع کدهای به کار رفته شامل کدهای اسکی می باشد.

همچنین چیز دیگری که شما به آن نیاز دارید کدهای Entity escape هستند که جایگزین برخی از کارکترها می شوند. می توانید جدول زیر را مشاهده نمایید:

Character		Escape Code
Ampersand	&	&
Single Quote	'	'
Double Quote	"	"
Greater Than	>	>
Less Than	<	<

کارکترهای escape در آدرس سایت جایگزین کارکترهایی که در بالا نشان داده شده اند می شوند. به آدرس زیر توجه کنید:

<http://www.yoursite.com/test.html&q=name>

این آدرس با استفاده از کدهای escape به صورت زیر در می آید:

<http://www.yoursite.com/test.html&q=name>

البته شما برای ایجاد نقشه سایت خود باید با تگ های XML که در نقشه سایت بکار می رود آشنا باشید. در زیر درباره این تگ ها توضیح داده ام:



1. فایل sitemap حتما باید با تگ آغازین `<urlset>` شروع شده و با تگ پایانی `</urlset>` خاتمه یابد.
2. تگ `<url>` به عنوان تگ منشاء (parent) برای هر آدرس معرفی شده استفاده گردد.
3. برای معرفی هر آدرس از تگ `<loc>` استفاده می شود.
4. `lastmod` برای اعلام آخرین تاریخ معرفی
5. `changefreq` مشخص کننده میزان و دوره تغییر می باشد.
6. `priority` میزان اهمیت صفحه را مشخص می کند که به صورت پیشفرض 0.5 می باشد.

توضیح تگ های مورد استفاده

`<urlset>` (اجباری)

مشخص کننده نوع استاندارد مورد استفاده در فایل xml

`<url>` (اجباری)

تگ منشاء برای آدرس های معرفی شده. تگهای بعدی در بین دو تگ ابتدایی و انتهایی این تگ قرار می گیرند.

`<loc>` (اجباری)

مشخص کننده آدرس صفحه معرفی شده. آدرس وارد شده در این قسمت، می بایست به صورت کامل وارد شود و نباید بیش

از 2048 کاراکتر باشد.

`Lastmod` (اختیاری)

مشخص کننده آخرین تغییر یا اصلاح می باشد

`Changefreq` (اختیاری)

این یک دستور نیست و تنها به عنوان یک تذکر به کار گرفته می شود

`Priority`

میزان اهمیت صفحات را مشخص می کند م و مقدار متغیری بین 0 تا 1 دارد.

البته باید بگم که تعیین میزان اهمیت صفحات تاثیری در جایگاه آدرس صفحه شما در موتورهای جستجو ندارد.

می توانید از نمونه زیر الگو بگیرید:

```
<?xml version="1.0" encoding="UTF-8"?>
<urlset xmlns="http://www.sitemaps.org/schemas/sitemap/0.9">
<url>
<loc>http://www.yoursite.com/</loc>
<lastmod>2010-01-01</lastmod>
<changefreq>monthly</changefreq>
<priority>0.8</priority>
</url>
<url>
```



```
<loc>http://www.yoursite.com/test?item=12&desc=vacation\_hawaii</loc>
<changefreq>weekly</changefreq>
</url>
<url>
<loc>http://www.yoursite.com/test?item=73&desc=vacation\_new\_zealand</loc>
<lastmod>2008-12-23</lastmod>
<changefreq>weekly</changefreq>
</url>
<url>
<loc>http://www.yoursite.com/test?item=74&desc=vacation\_newfoundland</loc>
<lastmod>2004-12-23T18:00:15+00:00</lastmod>
<priority>0.3</priority>
</url>
<url>
<loc>http://www.yoursite.com/test?item=83&desc=vacation\_usa</loc>
<lastmod>2004-11-23</lastmod>
</url>
</urlset>
```

اگر نقشه سایت شما در آدرس

<http://www.yoursite.com/test/sitemap.xml>

قرار داشته باشد، هر آدرسی که با <http://www.yoursite.com/test> شروع می شود می تواند در نقشه سایت شما قرار بگیرد، و اگر آدرسی مثل

<http://www.yoursite.com/new>

داشته باشد نمی تواند در نقشه سایت شما قرار گیرد. برای حل این مشکل ما به شما سفارش می کنیم هنگام ایجاد نقشه سایت از آدرس پایه سایت خود یعنی <http://www.yoursite.com> برای ایجاد نقشه سایت خود استفاده نمایید. شما می توانید نقشه سایت خود را به فایل robots.txt اضافه نمایید. این نکته هم عرض کنم که تفاوتی ندارد در چه قسمتی از فایل robots.txt نقشه سایت را قرار می دهید. چیزی که مهم می باشد این است که برای معرفی نقشه سایت خود در فایل robots.txt از دستور `<sitemap_location>` استفاده نمایید. به جای location باید آدرس محلی که نقشه سایت شما در آن جا قرار دارد، جایگزین نمایید به مثال زیر توجه کنید:

Sitemap: `<sitemap http://www.yoursite.com/sitemap.xml>`

شما می توانید از بیش از یک نقشه سایت استفاده کنید برای این منظور باید از sitemapindex استفاده کنید. که این کار معمولاً رایج نیست ولی شما می توانید با مراجعه به سایت ما نحوه ایجاد چندین نقشه سایت را فرا بگیرید. درباره محدودیت یک نقشه سایت هم خدمت شما عرض کنم که نباید بیشتر از 50000 آدرس و فضایی معادل 10MB برای هر نقشه سایت داشته باشید. هم چنین اگر از چند نقشه سایت در یک site-map index استفاده می کنید نباید بیش از 1000 نقشه سایت استفاده کنید.

ثبت نقشه سایت

شما علاوه بر قرار دادن نقشه سایت خود در فایل robots.txt می توانید آن ها را در موتورهای جستجو مانند گوگل ثبت نمایید. البته با ثبت نقشه سایت خود این انتظار را نداشته باشید که تمام صفحات شما crawl می شود. اما همانطور که قبلا عرض کردم باز سفارش می کنم که حتما نقشه سایت خود را ایجاد کرده و ثبت نمایید چون با این کار می توانید Crawlerها را برای بررسی صفحات خود راهنمایی نمایید.

SEO Spam و طراحان صفحات وب

تا حالا برای شما هم پیش اومده است که در حین جستجو در گوگل به سایت های برخورد کنید که بعد از وارد شدن به این سایت ها، فاقد محتویات مورد نظر هستند. در واقع این جور سایت ها با فریب دادن موتورهای جستجو در صفحات اول موتورهای جستجو یافت می شوند که به جزء ضرر به کاربر هیچ سودی دیگری ندارند. هدف این جور سایت ها، افزایش ترافیک سایت، افزایش رتبه السکا، افزایش پیج رنک، خراب کردن دنیای مجازی می باشد که این تعداد این گونه سایت ها در کشور ما متأسفانه سر به فلک می کشد.

بسیاری از فعالیت هایی که برای بهینه سازی سایتها و پایگاههای الکترونیک مورد استفاده قرار می گیرد، مورد علاقه و تایید موتورهای جستجو است زیرا ماهیت این فعالیت ها، بهبود وضع سایت ها برای استفاده و ورود رباتهای موتور جستجو است. اما رده خاصی از بهینه سازی نیز باعنوان بهینه سازی Black Hat یا روشهای تقلب وجود دارد که با واکنش شدید موتورهای جستجو مواجه می شود. این فعالیت ها می تواند تنوع و انواع زیادی داشته باشد با استفاده از مولفه های مورد توجه موتورهای جستجو تأثیری را بر نحوه رتبه بندی نهایی داشته باشد اما به هیچ عنوان استفاده و نفعی برای کاربر نهایی نداشته و در صورت آگاهی سیستم های جستجو به شدت با آنها مقابله شده و حتی ممکن است به حذف سایت از بانک داده ها منجر شود و ورود مجدد سایت نیز به مجموعه غدقن گردد. نکته مهم در روشهای تقلب این است که حتی ممکن است در یک بازه زمانی موفقیتی در کسب رتبه های بالا بدست آید اما کوتاه و بسیار گذرا خواهد بود.

حال از کجا ما بفهمیم عملیاتی که ما در پروسه سئو خود انجام می دهیم درست می باشد یا اشتباه است؟

من پس از مطالعه و تست کردن تا به حال حدود 34 مورد که نویسندگان و طراحان صفحات وب در زمینه SEO Spam به کار می برند را بررسی نموده ام. باید خدمت شما عرض کنم که شخصا قادر بر این نیستم که همه ی موارد را خدمت شما عرض کنم به دلایل مختلف که شاید مهمترین آن این باشد که بعضی از این موارد به صورت نجومی رتبه شما را ارتقا می دهند و حتی بعضی از آن ها توسط موتورهای جستجو هنوز قابل تشخیص نمی باشند ولی به هر حال در دسته ی اسپم ها قرار می گیرند. من در زیر تعدادی از این موارد را که شناخته شده هستند و متأسفانه به صورت رایج در صفحات وب مشاهده می کنیم خدمت شما عرض می نمایم. قبلا از آن اجازه بدهید تجربه ی و دانش خودم را در این زمینه در اختیار شما قرار داده و یک سری نکات را به شما عرض نمایم:

اقداماتی را که شما قبلا تجربه آن را نداشتید و مطمئن نیستید این اقدام درست یا نادرست را انجام ندهید:



تمام اقداماتی را که این نگرانی را برای شما ایجاد می کند که احتمال دارد سایت شما توسط موتورهای جستجو جریمه شود را به هیچ وجه انجام ندهید. این اقدامات می تواند شامل استفاده از متن های مخفی و ... باشد.

در معرفی وبسایت خود نباید به گونه ای عمل کنید که :

در عمل سایت شما آن گونه نباشد. به سادگی می توان کاری انجام داد که چیزی که بی ارزش است جذاب نمایش داده شود. خیلی از مدیران و فروشندگان ناموفق خیلی اوقات اینگونه عمل می نمایند اما بیاد داشته باشید استفاده از این استراتژی نادرست ممکن نابودی تجارت و همچنین عقب ماندن از رقبای حرفه ای خود می شوید. اتخاذ ساختار نا مناسبی برای لینک های سایت خود و اغراق گویی و دروغ گفتن درباره ی محصولات و خدمات خود در این زمینه از این دسته موارد می باشد. که این عمل علاوه بر اینکه موجب از دست دادن مشتریان و کاربران شما می شود، موجب بی اعتمادی موتورهای جستجو به سایت شما نیز می شود.

عدم اعتماد به افرادی که به اصطلاح خود را متخصص سئو معرفی کرده و شما را تشویق به استفاده از آیتم های غیر مجاز می کنند:

خیلی افراد معتقدند که اگر SEO Spam از لحاظ فنی نادرست باشد شما می توانید از آن در راه درست و صحیح استفاده نمایید ولی باید در جواب چنین افرادی اصطلاح معروفی که ما در کارمان داریم را یاد آور شوم که Spam is spam. چیزی که بد و نادرست است از آن نمی توان در یک کار تخصصی و اخلاقی استفاده نمود. حتی اگر شما این کار را انجام داده و در مقطعی هم موفق باشید مطمئناً باشید از دست موتورهای جستجو در امان نخواهید بود.

SEO spam به نام spamdexing نیز نامیده می شود چون شما به گونه ای از SEO spam، در ایندکس شدن صفحات خود در موتورهای جستجو استفاده می نمایید. عمل SEO spam در شکل های متفاوتی صورت می گیرد. بعضی از این تکنیک ها آشکار و قابل تشخیص و بعضی از آن ها مخفی و غیر قابل تشخیص می باشند. نکته ای که من باید خدمت شما عرض کنم این است که تمام موارد قبلاً تست شده و در خیلی از آیتم ها نیز جواب گرفته شده است ولی به این مطلب توجه داشته باشید که ممکن است با استفاده از SEO spam یک نتیجه خوب اولیه بگیرید ولی چیزی که مسلم است با هوشمندتر شدن و همچنین تغییر الگوریتم موتورهای جستجو روز به روز میدان برای ارز اندام SEO spam ها تنگ تر شده و ممکن است پس از چند ماه تا چند سال سایتی را که خیلی برای آن زحمت کشیده اید را از دست رفته ببینید و توسط موتورهای جستجو جریمه شوید :

Transparent links:

به آن دسته از لینک ها گفته می شود که در صفحات شما قرار دارد ولی کاربران نمی توانند آن ها را ببینند چون با پس زمینه صفحات شما دارای رنگ یکسانی می باشند.

Hidden links:

این دسته از لینک ها نیز در صفحات شما قرار دارد اما در میان عناصر سایت شما از جمله تصاویر صفحات مخفی شده اند. این لینکها تصاویر را به یک hyperlink تبدیل نمی کنند ولی موتورهای جستجو آن ها را پیدا کرده و دنبال می نمایند حتی اگر برای کاربران قابل تشخیص نباشد

Misleading links

لینک های فریب دهنده آن دسته از لینک ها می باشند که به ظاهر شما را به بخشی راهنمایی می نمایند که آن بخش با لینک داده شده به هیچ وجه ارتباطی ندارد.

Keyword stuffing

این مورد شامل به کارگیری کلمات کلیدی به صورت مکرر و پشت سر هم به صورت عمدی در متن ها و متا تگ ها می باشد.

Meta tag stuffing

این مورد هم مثل مورد بالا می باشد یعنی به کارگیری مکرر و پشت سر هم کلمات کلیدی در متا تگ ها می باشد.

Scrapper sites

این گونه وب سایت ها از آن دسته از سایت هایی می باشند که صفحات خود را با کپی برداری از دیگر صفحات موجود در وب پر می کنند . موتورهای جستجو به شدت از این گونه سایت ها متنفر می باشند زیرا از دید آن ها اینگونه سایت ها ، سایت های اصل و دارای اعتبار نمی باشند . بعضی از این گونه سایت ها اغلب با استفاده از Phishing یا سایر جرایم اینترنتی به دنبال سرقت اطلاعات شخصی مانند اطلاعات کارت های اعتباری و بانکی کاربران می باشد.

Links in punctuation

همان طور که گفته می شود هکر ها انسان های باهوشی هستند ، من باید خدمت شما عرض نمایم که افرادی هم که Seo spam را به کار می برند افراد خیلی باهوشی هستند چون در بازی ای ، مثل یک بازی شطرنج با موتورهای جستجو رقابت می کنند . و چرا این امر نیازمند هوش بالاست ، به این دلیل که موتورهای جستجو شناخته شده دارای قاعده، اصول و رفتار ثابت و مشخصی نیستند. Links in punctuation یک عمل هوشمندانه است. به قطعه کد زیر توجه کنید:

 Include real words here



.

اگر یک طراح سایت باشید با یک نگاه متوجه می شوید که منظور از کد دوم چیست، بله در کد دوم با کلیک بر روی یک نقطه وارد لینک مورد نظر می شوید، حال نکته این کد کجاست؟ با توجه به اینکه ما به یک نقطه لینک داده ایم کاربر توجه ای به این نقطه ندارد که بخواهد بر روی آن نقطه کلیک نماید ولی همین نقطه توسط موتورهای جستجو قابل تشخیص می باشد.

Hidden text

این گونه متن ها دارای رنگ یکسانی با رنگ پس زمینه هستند که برای کاربران قابل دیدن و تشخیص نمی باشند، ولی برای موتورهای جستجو قابل تشخیص هستند و همچنین معمولا این گونه متن ها ربطی با آن سایت پیدا نمی کنند و صرفا جهت فریب موتورهای جستجو مورد استفاده قرار می گیرند

Duplicate content

مطالب تکراری در وب سایت ، این اقدام فریبکارانه بدین منظور انجام می شوند که به موتورهای جستجو القا گردد که سایت مورد نظر با کلمات کلیدی خود دارای ارتباط زیادی می باشد و سایت خود را به صورت غیر از چیزی که در حال حاضر می باشد نشان دهد، باید خدمت شما عرض کنم که با توجه به اینکه موتورهای جستجو نسبت به این موضوع خیلی حساس می باشند ، افرادی که این کار را انجام می دهند سند نابودی سایت خود را امضا نموده اند.

Link-only pages

به صفحاتی گفته می شود که صرفا شامل لینک تنها می باشند که باید از این گونه صفحات خودداری کنیم، ولی یک استثنا هم در این زمینه وجود دارد و آن هم نقشه سایت می باشد که برای اینکه کاربران بتوانند به سرعت و راحت به مطالب سایت دسترسی پیدا کنند مورد استفاده قرار می گیرید

Page hijacking

به عملی گفته می شود که با کپی کردن صفحات وبی که مشهور و پربازدیدند و در اینترنت وجود دارند همراه می باشد و برای نمایش دادن سایت خود در موتورهای جستجو مورد استفاده قرار می گیرد ، وقتی کاربر لینک شما را در نتایج موتورهای جستجو مشاهده می کند با توجه به اینکه متن و موضوع جذاب و پربازدید می باشد بر روی لینک کلیک کرده و وارد سایت شما می شود ، در حالی که ماهیت سایت شما با مطلب مشهور قرار داده شده در سایت شما تناسب ندارد، در این زمینه باید عرض کنم که اکثر فروشگاه های الکترونیکی (که فروشگاه الکترونیکی نمی باشند) در ایران بدین شکل هستند، بدین صورت



که در کنار فروش محصولات خود از مطالبی که مثلا در میان قشر جوان پرتعداد است استفاده می نمایند و بدینگونه دست به جذب کاربر و ارتقای سایت خود در نتایج موتورهای جستجو می زنند.

آیا ساب دامین برای سئو سایت خوب است یا بد است؟

در پاسخ باید خدمت شما عرض کنم که ساب دامین ها می توانند به سئو سایت کمک نمایند به شرطی که به درستی از آن ها استفاده شود. استفاده درست از ساب دامین ها بدین منظور است که سایت شما با انتظارات کاربران مطابقت داشته باشد و به همان گونه ای باشد که به موتورهای جستجو معرفی شده اند. در واقع ساب دامین ها ارتباط مستقیمی با آدرس اصلی سایت شما دارند و با جایگاه سایت شما در موتورهای جستجو ارتباط مستقیمی دارند. یعنی هم می توانند موجب ارتقا رتبه سایت شما شوند و هم می توانند موجب جریمه شدن سایت شما گردند.

رابطه دامین و هاستینگ (لینوکس یا ویندوز) با سئو سایت :

طول عمر دامین تاثیر بسزایی دارد. در این زمینه علت های متفاوتی وجود دارد ولی چیزی که مشخص است که دامین های با طول عمر زیاد این تعهد و اعتبار و اطمینان را به موتورهای جستجو می دهد که مورد بررسی دائمی و دقیق قرار بگیرند و همچنین یکی دیگر از دلایل تاثیرگذاری طول عمر دامین بر سایت شما، سو استفاده هایی است که توسط هکرها صورت می گیرد. معمولا هکرها از دامین هایی برای مقاصد خود استفاده می نمایند که دارای طول عمر کمتر از یکسال می باشند. سعی کنید هنگام انتخاب دامین ، دامینی کوتاه انتخاب نمایید و از انتخاب دامین طولانی به غیر از در حالت ضروری و خاص بپرهیزید. چون دامین کوتاه و ساده در ذهن کاربر خواهد ماند ولی دامین طولانی بدین صورت نیست.

ما دونهوع هاست اصلی و شناخته شده داریم، Linux و Microsoft Windows. که شما با توجه به هزینه ای که خواهید نمود و همچنین تکنولوژی هایی که در طراحی سایت خود به کار می برید می توانید یکی از آن ها را انتخاب نمایید. معمولا سرورهای لینوکس نسبت به سرورهای ویندوز ارزاتر خدمات می دهند در ضمن خدماتی هم که ارائه می دهند دارای امکاناتی زیادی از لحاظ اکانت زیردامنه - ایمیل - اکانت FTP و ... هستند. سرورهای لینوکس بنا به ویژگی های سیستم عامل حاکم بر آنان دارای ثبات و امنیت بیشتری نسبت به سرورهای ویندوز هستند و کمتر در معرض تهدید قرار می گیرند. چون اکثر سرورها از سخت افزارهایی در یک سطح استفاده می کنند سرعت آنها چندان با یکدیگر تفاوتی ندارند ولی برخی از کاربران خدمات میزبانی معتقدند که سرورهای لینوکس سریعتر از سرورهای ویندوز به بازدید کنندگان سرویس ارائه می دهند. سرورهای ویندوز بیشتر از سرورهای لینوکس از نرم افزار های برنامه نویسی و ایجاد بانک های اطلاعاتی پشتیبانی می کنند؛ ولی این پشتیبانی مخصوصا در مورد PHP به دلیل ناقص بودن و امنیت پایین؛ کاربرد محدود دارد. به گونه ای که توصیه میشود از هاستینگ ویندوز برای سایت های طراحی شده با php/mysql نباید استفاده نمود.



و اما نکته که شما باید به آن خیلی توجه کنید این است که پلتفرم ویندوز و لینوکس در زمینه سئو یک تفاوت خیلی مهم دارند. و آن هم این است که نام فایل های ویندوز حساس به بزرگی و کوچکی حروف نیستند و این در حالی است که نام فایل های لینوکس نسبت به بزرگی و کوچکی حروف حساس هستند. اجازه بدید یک مثال برای شما بزنم :

شما در یک سیستم ویندوز نمی توانید دو فایل `a.html` و `A.html` را در یک دایرکتوری یا فلدر قرار داده و داشته باشید. این در حالی است که شما می توانید در سیستم لینوکس فایل های `a.html`, `A.html`, `a.HTML` و `a.HTML` در یک فلدر قرار دهید و هر کدام از این فایل ها یک فایل جداگانه محسوب شود.

البته اختلاف زیادی برای انتخاب هاست بهتر برای سئو وجود دارد ولی مسئله ای که اکثرا با آن موافق هستند و به ویژه در سایت های جوملایی باید به کار گرفته شود هاست لینوکس می باشد.

با تشکر