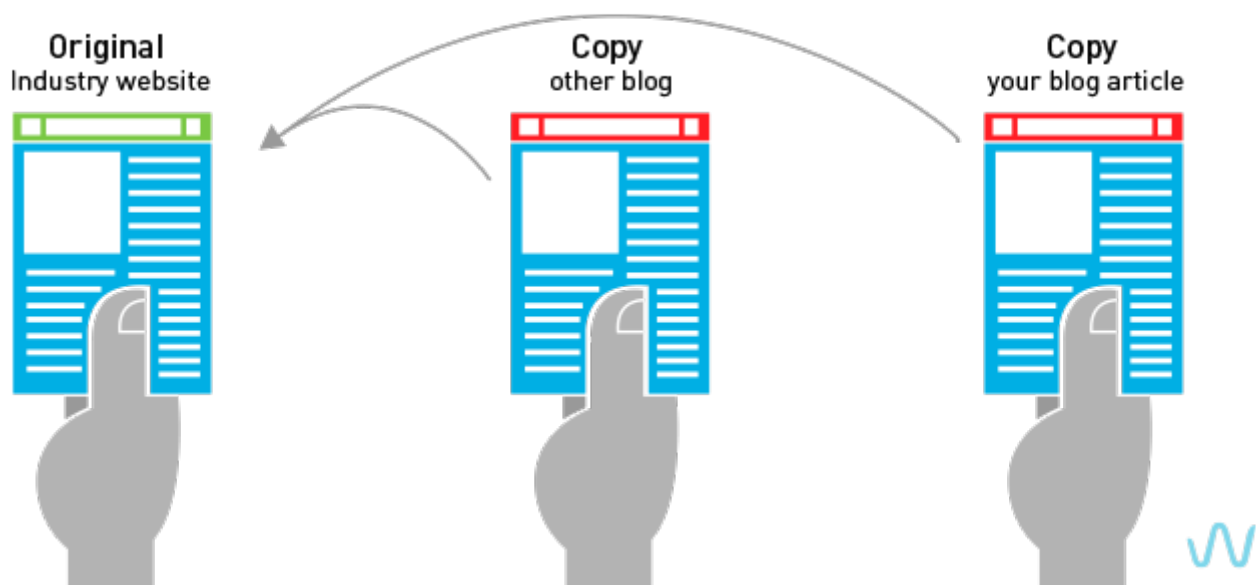


محتوای تکراری در طراحی سایت چیست و انواع روش‌های رفع مشکل محتوای تکراری کدامند؟



امروزه علمی به نام سئو (Search Engine Optimization یا seo) نقشی کلیدی در موفقیت یک سایت یا وبلاگ بازی می‌کند، البته این علم چیزی نیست جزء رعایت یک سری موارد در جهت ایجاد محتوای استاندارد در وب و آگاهی از نحوه عملکرد موتورهای جستجوگر به عنوان رکن تعیین کننده در افزایش بازدیدها و از طرفی ایجاد محتوای مفید و کاربرپسند برای یاری رساندن به مخاطبان و در نتیجه جلب محبوبیت بیشتر برای سایت یا وبلاگ شما نزد آنان، اما آنچه که سبب می‌شود سئو را علمی پیچیده در وب لقب دهند بیشتر به خاطر گستردگی فاکتورهایی است که در آن نقش بازی می‌کنند، یعنی این علم خود از زیرشاخه‌ها و نکاتی فرعی تشکیل شده که رعایت آنها به صورت یک مجموعه در کنار هم می‌تواند نقش موثری در موفقیت شما در عرصه وب داشته باشد؛ یکی از این فاکتورها داشتن محتوای یکتا و پرهیز از ایجاد عمدی یا سهوی صفحات و مطالب تکراری (duplicate content) است. از آنجا که موتورهای جستجو جهت خزیدن و تجزیه و تحلیل این صفحات زمان صرف مینمایند لذا در کیفیت نتایج جستجو نیز تأثیر می‌گذارد.

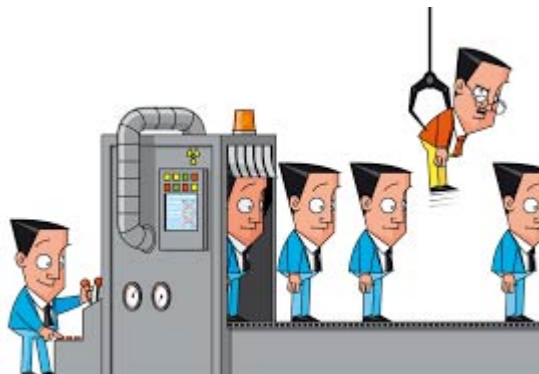
محتوای تکراری چیست؟

محتوای تکراری تعاریف مختلفی دارد که در زیر به سه مورد از آن اشاره شده است:

تعریف اول: در برخی مواقع نیز عنوان و شرح یک صفحه مشابه نیست اما متون آن یکسان است.

تعریف دوم: گاهی کل محتوای یک صفحه با صفحه همسان دیگری مشابه است.

تعریف سوم: گاهی تگ عنوان و شرح یک صفحه با صفحه دیگر تشابه دارد.



ریشه اصلی مشکل محتوای تکراری چیست؟

محتوای تکراری از ارسال مطالب مشابه در صفحات و یا وب سایتهای مختلف وجود دو یا چند دامنه برای یک سایت و همچنین از برنامه نویسی توسعه نیافته با قوانین سئو ایجاد میگردد.

مورد اول : خیلی جالب و فنی است، و ریشه اصلی آن HTTP protocol میباشد زیرا نمیتواند یک روش استاندارد برای شناخت بهترین URL یک صفحه را معرفی نماید. یعنی اینکه به یک صفحه با آدرسهای متفاوت دسترسی داشته باشیم و در عین حال هیچ اطلاعاتی در رابطه با URL مرجع وجود ندارد.

مورد دوم :برخی از مدیران سایتها برای گسترش وب سایت خود محتوا، مقالات و سایر اطلاعات را در سایتهای دیگر قرار میدهند. به عبارت دیگر از استراتژیهای نادرست لینک بلدینگ جهت افزایش back link استفاده مینمایند. بنابراین در این مورد ریشه اصلی مشکل کسانی هستند که با استفاده از تکنیکهایی که با نام کلاه سیاه سئو شناخته میشود، سعی در گسترش سایت دارند.

مورد سوم : توسط شرکتهای که تمایل دارند چند دامنه داشته باشند به وجود میآید. به عنوان مثال example.com و example.co.uk دو دامین میباشد که هر دو آنها یک محتوا را نمایش میدهد اما به دلیل عدم استفاده از ۳۰۱ Redirect مشکل محتوای تکراری ایجاد میشود. بنابراین در این مورد نیز صاحبان وب سایتهای و شرکتهای طراحی سایت عامل اصلی ایجاد محتوای تکراری هستند، کسانی که نمیدانند چگونه از ۳۰۱ جهت ست کردن دامین اصلی و فرعی استفاده نمایند.

به عنوان مثال:

<http://www.example.com/index.html?somevar>

<http://www.example.com/index.html>

<http://www.example.com>

<http://www.example.com/index.htm>

<http://www.Example.com>

<http://example.com>

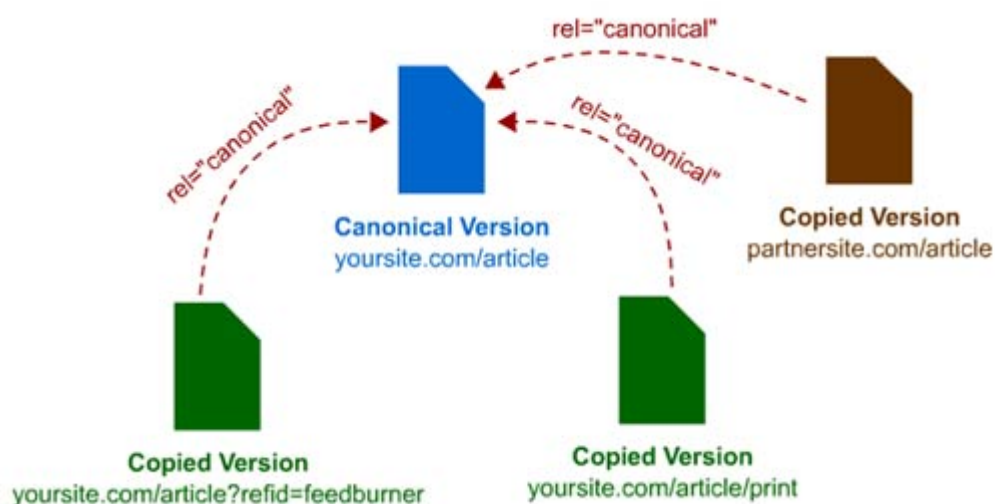
</http://www.example.com>

فناوریها و زبانهای متفاوتی (Asp.Net, PHP) برای پیاده سازی وب سایتهای داینامیک وجود دارد و توجه به اینکه تکنولوژیهای متفاوت ویژگیهای متفاوتی را ساپورت مینمایند وضعیت را کاملاً پیچیده میسازد. همه مشکلات بالا برای کسانی که نمیدانند موتورهای جستجو چگونه کار میکنند و چگونه از متدهای استاندارد برای پیاده سازی سایت استفاده نمایند، پیش میآید. تمام URL های بالا میتوانند به یک صفحه هدایت شوند، اما HTTP protocol میبایست یکی را به عنوان بهترین آدرس انتخاب و بقیه آدرسها را به همان صفحه هدایت نماید به عبارت دیگر <http://www.example.com> و <http://www.example.com/index.html> هم میتواند طوری پیاده سازی شوند که به یک صفحه هدایت شوند و یا دو صفحه کاملاً مجزا داشته باشند.



چرا محتوای تکراری یک مشکل است؟

برای درک راحت‌تر محتوای تکراری که چه مشکلاتی را درست میکند، را از نقطه نظر موتورهای جستجو مختلف بررسی می‌کنیم. به طور مثال موتورهای جستجو ابتدا در صفحات سایت می‌چرخند، پس از آن نسبت به تجزیه و تحلیل آن اقدام مینمایند و زمانیکه تجزیه و تحلیل تمام شد، به هر یک از صفحات امتیازی میدهند. هنگامی که کاربر کلمه‌های را در گوگل یا هر موتور جستجوی دیگری وارد مینماید، ابتدا جدول شاخصهای خود را بررسی و سپس نتایج را به کاربر نمایش میدهد که هر چه این کار سریعتر انجام شود بهتر است. زمانیکه صفحاتی با محتوای تکراری داریم اولاً موتورهای جستجو زمان زیادی را برای ایندکس نمودن آنها صرف مینمایند و از طرفی هنگامیکه کاربران در نتایج جستجو صفحاتی با مطالب یکسانی ببینند به طور حتم، کلمه کلیدی خود را تغییر خواهند داد و یا کلاً از جستجو منصرف میشوند. بنابراین محتوای تکراری بر روی رضایت کاربران تأثیر میگذارد زیرا نتایج جستجوی با کیفیتی نداریم.



چگونه مشکل محتوای تکراری را برطرف نماییم؟

در حالت اول، در صورتیکه نیاز به ثبت چندین دامنه برای وب سایت خود دارید، ابتدا یک دامنه را به عنوان دامنه اصلی انتخاب و بقیه دامنه‌ها را با استفاده از HTTP ۳۰۱ Redirections تنظیم نمایید. به عنوان مثال در صورتیکه دو دامنه با عنوانهای (اصلی example.com و example.co.uk) داشته باشید. دامنه example.co.uk میبایست با استفاده از HTTP ۳۰۱ Redirections تنظیم شود. تا هر زمان که آن را وارد کریم به example.com هدایت شویم. روشهای مختلفی برای HTTP ۳۰۱ Redirections وجود دارد که از جمله آنها میتوان به تنظیمات PHP/ASP/JSP Redirection، htaccess، DNS و غیره اشاره نمود.

مورد دوم کمی پیچیده‌تر است و همان طور که اشاره شد، موتورهای جستجو با تنظیم قوانین سعی دارند که اثرات محتوای تکراری را به حداقل برسانند. آنها از مجموعه قوانین که برای URL تنظیم شده، برای انتخاب بهترین نسخه استفاده مینمایند. بنابراین پیشنهاد میکنیم

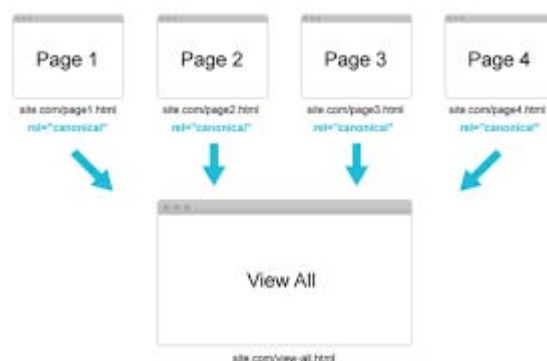
که حتماً ساختار لینکهای خود را بررسی نمایید. روی ساختار لینکهایتان کار کنید. کاری که شما باید انجام دهید این است که مطمئن شوید تمام لینکهای سایتتان به بهترین URL هدایت میشوند و موقعیتی وجود ندارد که دو لینک متفاوت به یک صفحه هدایت شوند. در مورد سوم، راه حل نسبتاً ساده است. فقط کافیست از ارسال محتوای یکسان به منابع متعدد اجتناب نمایید و همیشه مطمئن شوید که از تکنیکهای white hat سئو استفاده مینمایید. رعایت این نکات در گرفتن لینک بیشتر و از همه مهمتر افزایش ترافیک سایت تأثیر میگذارد.

مهمترین چیزهای که شما باید در نظر داشته باشید: تمام پارامترهای کم اهمیت و اضافی همانند (SESSIONIDS,Sorts,Specifies) و غیره) را از URL حذف نمایید. اگر وب سایت شما با درج عبارت WWW در نوار آدرس و در عین حال بدون درج آن نیز در دسترس کاربران قرار میگیرد، به طور حتم با مشکل مطالب تکراری روبرو خواهید شد، چرا که از دید موتورهای جستجو این دو آدرس، دو صفحه مجزا تلقی شده و به طور جداگانه ایندکس میشوند، البته برخی از آنها نظیر گوگل تا حدودی قادر به تشخیص آدرس پیش فرض هستند، اما بهتر است یک آدرس را برگزینید و حالت دیگر را با ارسال کد ۳۰۱ HTTP، ریدایرکت نمایید. دسترسی به سایت با و بدون عباراتی چون index.php، index.html، index.asp و غیره در ایجاد محتوای تکراری تأثیر گذار است؛ شاید گمان کنید موتورهای جستجو تشخیص میدهند که صفحه ایندکس یک سایت، همان صفحه اصلی آن است، اما در پاسخ باید بگوییم که معمولاً اینطور نیست، در برخی از سرورها ممکن است تنظیمات پیش فرض تغییر داده شود و صفحه اصلی سایت با عبارتی دیگر تنظیم شده باشد؛ لذا فایل ایندکس الزاماً به معنی صفحه اصلی نیست، به همین دلیل موتورهای جستجو ممکن است آن را به صورت جداگانه ایندکس نماید، لذا بهتر است صفحه دوم را به صورت استاندارد ریدایرکت نمود. URL هایتان را با استفاده از / به پایان برسانید. Canonical Tag در واقع آدرسی است که به صفحه مرجع خود اشاره میکند، این آدرس توسط تگ link در header صفحه تکراری قرار میگیرد (در تگ head) همانند زیر: این تگ توسط خصوصیت "rel="canonical" برای موتورهای جستجو متمایز میشود و زمانی که ربات موتور جستجو به صفحاتی که شامل این تگ هستند میرسد، با آن به عنوان یک صفحه کپی برخورد مینماید و کلیهی محتوا و رنگ آن را به حساب آدرس ذکر شده قرار میدهد.



چگونه محتوای تکراری ایجاد میشود؟

دلایل اشکار برای ایجاد این مشکل، کپی کردن یک متن در صفحه های مختلف با URL های متفاوت میباشد و دلیل دیگر آن به استفاده از تکنیکهای کهنه و قدیمی و برنامه نویسی ضعیف بدون در نظر گرفتن قوانین سئو در طراحی و پیاده سازی وب سایت بر میگردد.



آیا محتوای تکراری بر رتبه سایت در موتورهای جستجو تأثیر میگذارد؟

موتورهای جستجو همیشه در پی یافتن مطالبی یکتا از URL های شما هستند تا نتایجی بهتر به کاربران خود نشان دهند، وقتی چند صفحه با URL متفاوت و محتوای یکسان را معرفی میکنیم، ابتدا محتوای تکراری را شناسایی میکند، سپس سعی میکند که صفحه اورجینال را بیابد که به طور معمول این انتخاب را بر اساس عمر صفحه، رنک صفحه و خیلی از گزینههای دیگر انجام میدهد. گوگل در آخرین آپدیت خود با نام گوگل پنگوئن سعی نموده که صفحاتی با محتوای تکراری را به عنوان اسپم بشناسد و به آنها امتیاز منفی میدهد. تکرار مطالب یکسان و تکراری و URL تکراری ممکن است موجب شناسایی سایت یا وبلاگ شما به عنوان اسپم شود که در

نهایت موجب حذف آدرس و اطلاعات شما از لیست جستجو و پایگاه داده موتورهای جستجو میگردد. اگر صفحات تکراری در سایت های خود کاهش دهید در این صورت می توانید اعتماد موتورهای جستجو را جلب خواهید کرد که این کار بر روی رتبه سایت شما تأثیر میگذارد. از طرفی در صورتی که محتوای تکراری داشته باشد در نتایج جستجو پراکنده میشود و در نتیجه بر روی کاهش رنک صفحه تأثیر میگذارد. زیرا در حالت عادی بازدیدها به یک مطلب تعلق میگیرد و رنک آن نیز به همان نسبت افزایش مییابد و روشن است که اگر این رنک را بین چندین صفحه تقسیم کنیم در نهایت رنک تمام آنها کاهش مییابد.

