حل مسئله بودجه خزش، يكبار براى هميشه!

0

زمستان سرد

ىندكس

NDEX



ابوالفضل پارسائی



تشنگان یادگیری، دوستان من هستند

سلام میکنم به دوست ناشناسم که این کتاب جذاب و ارزشمند را برداشتهای تا بخوانی، لـذت ببری و یـاد بگیـری! گـاهی شـبهایی را تـا صـبح در حـال ایـدهپردازی و آمادهسازی مطالب کتاب بودم و خوشحالم که الان منتشر شدهاست و حرفهای دل من را میخوانی؛ **اگر کتاب را دوست داشتی، منتظر نظرات ارزشمند و سازنده شما هستم**. (مطمئن باش با خواندن نظرات و پیشنهادهایت، خستگی از تنم فرار میکند و دیگر یک دوست ناشناس نیستی و من یک دوست آشنای خوب دارم!) تمام تلاش خود را کردم تـا مطالب فنی را سـاده، سـرگرمکننده و مفید بنویسم. قول میدهم **در حین مطالعه خسته نشوی و گاهی لبخند شیرینی روی لبانت بنشیند**. متشکرم از رفیـق دوست داشـتنی و آموزگـاری متعهـد، **مهنـدس امـین اسـماعیلی** و

که این فرصت را در اختیار من گذاشتند و همیشه همراه و پاریرسان من و تمام

دانشجویان آکادمی وبسیما هستند.

ابوالفضل پارسائی

نویسنده، گَرافیست، صفحهآرا و ایدهیرداز

in linkedin.com/in/abolfazlparsaie

فهرست رباتها

ایندکس چیست و گوگل چطور این کار را انجام میدهد؟

بودجه خزش چیست؟ چقدر باید به آن اهمیت بدهیم؟

تعیین بودجه خزش چگونه است و چطور مصرف میشود؟

مفاهیم بسیار مهم و مربوط به بهینهسازی بودجه خزش

بررسی مهمترین و اصلیترین ابزار: گوگل سرچ کنسول

تکنیکهای حرفهای برای افزایش انفجاری بودجه خزش

بررسی تلههای خزنده در سایت و رهایی از آسیب آنها

















 $\mathbf{)}$

Ò



^{بخش۱} داستان گوگل و آدرسهای وب (URLs)

تمام داستان ایندکسها از یک آدرس شروع میشود؛ آدرسی که گوگل بتواند آن را پیدا کند و به آن دسترسی داشته باشد.

اگر سوال شما این است که «**گوگل چطور یک آدرس را مییابد و به آن دسترسی پیدا میکند؟**» از رایجترین روشهای یافتن آدرسها این ۴ مورد هستند:

- بک لینکھا
- نقشه سایت
- درخواست از گوگل
 - ربات پیشرفته

استفاده از بک لینک (Backlink): از بقیه میپرسم!

گوگل در حال حاضر صدها میلیارد صفحه وب را در سرورهای خود ذخیره دارد؛ حال اگر مالک یکی از این صفحات، سایتمان را معرفی کند و به آن لینک بدهـد، <mark>گوگـل بـا</mark> استفاده از این لینک متوجه سایتمان میشود</mark> و شروع به گشت و گذار در آن میکند و اگر چیز به درد بخوری را پیدا کند آن را برای خود ذخیره میکند!

استفاده از نقشه سایت (Site Map): خودم آدرس را بلدم!

نقشه سایت یک جور راهنما برای گوگل است که آن را ما میتوانیم به گوگل معرفی کنیم و به او بفهمانیم که **کدام صفحات و فایلهای سایت مهم هستند** و او با توجه به اولویتهای ما شروع به خزیدن در سایت کند!

نقشه سایت چیست؟ ساخت سایت مپ و معرفی به گوگ<u>ل</u>



米

درخواست مستقیم از گوگل (URL Inspection Tool): لوکیشن فرستادم!

خود **گوگل ابزاری را در سرچ کنسول برای ما قرار داده است تا با استفاده از آن بتوانیم صفحات خود را بـه او معرفی کنـیم** و از ایـن راه، رباتهـایش بتواننـد بـه صفحه مـورد نظرمان دسترسی داشته باشند و محتوای آن را بررسی کنند.

URL Inspection؛ سریعترین راه ایندکس دستی صفحه در گوگل

زیرکی گوگل: استفاده از رباتهای فوق پیشرفته و حرفهای!



9

✻

امین اسماعیلی صحبت شگفت انگیزی دارد: «**در دنیای امروز، همگی ما**

مهمترین و پیشرفتهترین رباتهای گوگل هستیم!»





آنقدر فضای نرم افزار گوگل کروم (Google Chrome) برای من بهینه است و به آن عادت کردهام که کمتر زمانی از مرورگر دیگری استفاده میکنم! بهنظر شما چرا گوگل ایـن میـزان سـرمایهگذاری را در توسـعه گوگـل کـروم انجـام داده و فضـای آن را بـه بهینهترین شکل ممکن طراحی کرده است؟

پاسخ ساده است! استخدام رباتهای فوق پیشرفته بهصورت رایگان! این مورد فقط مربوط به استفاده از مرورگر نیست؛ برای مثال <mark>گوگل فرایندهای نحوه جستجو، بررسی</mark> و انتخاب نتایج، واکنشهای بعد از آن و هزاران اطلاعات دیگر را از ما دریافت میکند! بنابراین **استفاده از دادههای گوگل کروم نیز یکی از راههای دسترسی به صفحات وب** ا**ست** و اگر ما با این نرم افزار وارد صفحهای بشویم، گوگل از وجود آن مطلع میشود. تا اینجا ۴ روش متداول یافتن آدرسها را بررسی کردیم در ادامه باید سراغ ۲ فرایند

دیگر یعنی خزش و ایندکس برویم.

بخش **کزش (Crawling): دور دور رباتها در اینترنت**

به فرایند بررسی و دانلود محتوا، ظاهر، کدها و هرچیزی که در یک صفحه است اصطلاحا خزش (Crawling) میگوییم.

گوگل، رباتهای کامپیوتری (Googlebot) بهنام عنکبوت (Spider) را طراحی کرده است تا وظیفه مهم خزش (Crawling) را انجام دهند. این رباتها با استفاده از آدرسها (URLs) به راه میافتند و ماجراجویی خود را در صفحات اینترنت ادامه میدهند. آنها هم مانند بسیاری از ما ایرانیها چند اسمه هستند! مثلا به آنها خزنده (Crawler) نیز میگوییم.

البته همیشه این اتفاق به این سادگی نیست! گاهی به این فرایند مرحله پردازش (rendering) نیز اضافه میشود.

بخش المنافع بخش بخش بخش بخش الما متوجه نمى شوم!

برای توضیح این فرایند باید یک مثال قابل درک بزنیم تا به راحتی مفهوم آن را متوجه شوید؛ پس خوب دقت کنید:

- سایتهای دنیای وب، مانند کشورهای دنیای واقعی هستند؛ هر کشوری ظاهر و زبان متفاوتی نسبت به دیگری دارد.
- رباتهای گوگل شبیه به گردشگرانی هستند که به کشورهای مختلفی سفر میکنند.
- زبان برنامه نویسی پایتون مثل زبان انگیسی یک زبان بین المللی است و گردشگران با تسلط به این زبان، در طول سفر مشکلی برای برقراری ارتباط نخواهند داشت.
- مرورگر کروم مانند یک مترجم است که به گردشگر کمک میکند تا درک درستی
 از گفتگوها داشته باشد.

حال فرض کنید ما یک گردشگر هستیم و فقط به زبان انگلیسی (زبان بینالمللی) مسلط هستیم و در سفر به دور دنیا وارد لبنان میشویم. همه افراد در این کشور در حال صحبت با زبان عربی هستند و ما چیزی از صحبتهایشان را نمیفهمیم یا فقط خیلی کم و محدود متوجه آن میشویم. برای اینکه بدانیم افراد چه میگویند و بتوانم با آنها ارتباط برقرار کنیم، یک مترجم عرب زبان را استخدام میکنیم تا گفتگوهای آنها را با زبان انگلیسی برای ما شرح دهد. ما نیز با استفاده از این ترجمه، به راحتی متوجه مکالمات میشویم و صحبتهای آنها را درک میکنیم.

تقریبا در فرایند پردازش و رندرینگ نیز چنین اتفاقی رخ میدهد.

اگر کراولر با صفحهای مواجه شود که از تکنولوژیهای جدید استفاده کند (مثلا در طراحی آن از یک زبان برنامه نویسی اختصاصی استفاده شده باشد)، <mark>کراولر خروجی</mark> خیلی کمی دریافت میکند یا با اخطار «شما باید جاوا اسکریپت داشته باشید» مواجه میشود. این خطا یعنی شما باید یک مترجم داشته باشید تا صفحه ما را درک کنید؛ در این مرحله، کراولر از مرورگر کروم استفاده میکند تا صفحه را برایش کامل باز (Load) و سپس اطلاعات مورد نیاز خود را دریافت و ذخیره کند و بعدش هم برود پی کارش...

به تلاشهای رباتهای گوگل برای درک و استخراج اطلاعات و مفاهیم کلیدی یک صفحه، پردازش (Processing) میگوییم. برای انجام این کار، گوگل دست به کالبد شکافی میزند و صفحه را تشریح (Render) میکند.

بخش المجنب المنابع المنابع المالي المالي المالي بع بخش آكادمي!

آقای پشمک تا حالا حتی از یک محتوای آکادمی وبسیما هم استفاده نکرده است! اگر از آقای پشمک بپرسیم «بهترین مرجع ایرانی برای سئو را کجا پیدا کنیم؟» قطعا اولین اسـمی کـه بـه زبـان مـیآورد وبسـیما نخواهـد بـود؛ چـون آقـای پشـمک اصـلا نمیدانـد وبسیمایی نیز وجود دارد! فقط کافیست یکبار از محتواهای ما استفاده کند تا اولین و آخرین اسمی که پیشنهاد بدهد وبسیما باشد! (این داستان ناشی از تخیلات پریشان نویسنده بود... حتی آقای پشمک هم محتواهای وبسیما را دنبال میکند!!!

فهرست بندی یا ایندکس هم دقیقا مانند داستان ما است! **ایندکس زمانی اتفاق** میافتد که اطلاعات پردازش شده از صفحات، ذخیره شده و به فهرست جستجو (Search index) اضافه میشوند.



فهرست جستجو دقیقا همان جایی است که گوگل برای ارائه نتایج به ما از آن استفاده میکند. وقتی ما عبارتی را جستجو میکنیم، گوگل فهرست نتایج خود را بررسی میکند و از بین آنها بهترین نتایج را برای ما رتبهبندی میکند؛ در نتیجه **اگر صفحات ما در فهرست سرچ نباشند یعنی گوگل از وجود ما بیاطلاع است و هرگز در نتایج نیز نخواهیم بود.** (مثل آقای پشمک که آکادمی وبسیما را نمیشناخت)

تا اینجا با مفاهیم اصلی ایندکس سایت آشنا شدیم و دانستیم که در هر مرحله چه اتفاقهایی در حال رخ دادن است. از همه مهمتر متوجه شدیم که <mark>اگر بهدنبال گرفتن</mark> نتـایج بسـیار خـوب از گوگـل هسـتیم در اولـین قـدم بایـد بـهدنبال اینـدکس شـدن سایتمان باشیم.

اما سوالی که اینجا مطرح میشود و ما برای پاسخ دادن به این سوال، این ایبوک را آماده کردهایم این است: «**چه صفحاتی باید ایندکس شوند؟**» بهنظر شما ایندکس صفحات را باید به حال خود بگذاریم تا گوگل هرچه در سایت هست و نیست را ذخیره کند؟ در ادامه قصد داریم مفهوم «بودجه خزش» (Crwal Budget) را بررسی کنیم.







بخش ا اهمیت بودجه خزش: نشستهام به در نگاه می کنم...

اگر در زمینه سئو تجربه داشته باشید، حتما برایتان پیش آمده است که بخشی از محتوای یک صفحه را تغییر میدهید و یا آن را بهروز میکنید اما گوگل پس از چند هفته متوجه این تغییرات میشود و آن را کراول و ایندکس میکند.

یکی از مهمترین رازها در راه رسیدن به موفقیت در سئو این است که **گوگل در سریعترین زمان ممکن و تا حد امکان، صفحات مهم ما را پیدا و درک کند** و اگر صفحه جدیـدی منتشـر یـا محتـوای یکـی از صـفحات را بـهروز کـردیم، گوگـل بهسـرعت آن را ایندکس کند!



اگر با یک زبان سادهتر بخواهم برایتان بنویسم: «<mark>هرچقدر زودتر صفحات ما ایندکس</mark> شود، زودتر نتیجه خواهیم گرفت!» حالا فرض کنید تعداد صفحات سایت ما بیشتر از بودجهای است که گوگل برای ایندکس برای سایتمان در نظر گرفته است! در این صورت صفحههای بسیاری را در سایت خود داریم که گوگل از آنها اطلاع ندارد یا ایندکس آنها را به تاخیر انداخته است.

ش ۲ 🔵 تعریف بودجه خزش: همیشه قبل خواب ایندکس شو و راحت بخواب

گوگل هرچقدر هم رباتهای پرتلاش و زحمتکشی برای یافتن صفحات اینترنت داشته باشد، منابع کافی بـرای دانلـود و ذخیـره تمـام صـفحات را نـدارد! میلیاردهـا صـفحه در فضای اینترنت وجود دارند و هزینههای ایندکس صفحات آنقدر زیاد است که اگر گوگل بخواهد تمام آنها را داشته باشد، قطعا ورشکست خواهد شد؛ اگر هم کسی ایـن منابع را داشته باشد، منابع خود را با ذخیره اطلاعات بیارزش و بهدرد نخور هدر نخواهد داد!

در نتیجه گوگل محدودیتهایی را برای رباتهای خود اعمال کرده و مقدار مشخصی را برای کراولینگ هر سایت اختصاص داده است به آن بودجه خزش (Crwal Budget) میگویند.

نکته بسیار مهم:

هر چیزی که در سایت ما خزیده شود، لزوما ایندکس نمیشود!

بعد از اینکه رباتهای گوگل صفحهای را پیدا و اطلاعات آن را استخراج کردند، **گوگل** این اطلاعات را ارزیابی میکند تا متوجه شود آیا این صفحه ارزش ایندکس شدن را دارد یا نه؟ به همین دلیل است که بعضی از صفحات سایت ما کراول شدهاند اما گوگل آنها را ایندکس نکرده است.

بودجه خزیدن، میزان سرعت و تعداد صفحاتی است که رباتهای گوگل در یک بازه زمانی مشخص، سایت شما را کراول میکنند؛ دو عنصر اصلی این بودجه را تعیین میکنند:

- محدودیت ظرفیت خزیدن (Crawl Limit)
- میزان نیاز و ارزشمندی صفحه (Crawl demand)

میزان پاسخگویی سرور به درخواست ایندکس (Crawl limit / host load) رباتهای گوگل، بسیار خودخواه و مغرور هستند و دوست دارند بدون محدودیت در سایت ما بچرخند و منابع سرور ما را نوش جان کنند! گوگل نیز از این موضوع آگاه است و برای اینکه این رباتهای کنجکاو و هوس باز با درخواستهای متعدد و مکرر **تمام منابع را استفاده نکنند** و باعث از کار افتادن (Down) سایت نشوند محدودیتهایی را برای رباتهای خود در استفاده از منابع سایت اعمال کرده است. بنابراین هرچقدر منابع سرور ما محدودتر باشد، رباتهای گوگل در کراول سایت

حال گوگل چطور متوجه میزان توان سرور ما میشود؟

۱. با استفاده از تجربه:

اگر رباتها هنگام خزش چندین بار با خطای سرور روبرو شده باشند و سرور توانایی پاسخ به آن را نداشته باشد، متوجه کمبود منابع سرور میشوند و **محدودیت بیشتری** را درنظر میگیرند. اگر سایت برای مدتی با سرعت بالایی به درخواستها پاسخ دهد این محدودیتها کاهش مییابد و **سرعت ایندکس سایت نیز بیشتر میشود**.

۲. بررسی تعداد سایتهای فعالی که روی هاست اجرا میشوند:

مثلا سایتهای بزرگی که از هاستهای اشتراکی استفاده میکنند که صدها سایت دیگر نیز در آن فعال باشند، **محدودیت کراول زیادی را تجربه میکنند.**

میزان نیاز و ارزشمندی صفحه (Crawl demand / crawl scheduling)

گوگل با استفاده از معیارهای زیر، تعیین میکند که **چه صفحههایی ارزش ایندکس شدن یا دوباره ایندکس شدن را دارند** میشوند:

۱. میزان اهمیت صفحه را اندازه گیری میکند:

چه تعداد لینکهای داخلی و خارجی باکیفیت به صفحه داده شده است و در حال حاضر این صفحه در عبارتهای کلیدی مربوطه چه رتبهای دارد؟

۲. با چه نوع صفحهای سر و کار دارد:

مثلا جنس و شرایط محتوایی صفحه درباره ما با صفحه دستهبندی متفاوت است. احتمال تغییر در کدام صفحه بیشتر است؟

۳. تازگی محتوا:

آیا محتوای صفحه بهروزرسانی میشود؟ بازه زمانی بهروزرسانی آن چند وقت یک بار است؟

داستان گل مراد، تاجر بزرگ: باشد تا رستگار شوید...

خب بچههای عزیز یکی بود یکی نبود!!! در زمانهای قدیم مرد فقیری بهنام «**گل مراد**» زندگی میکرد که ثروت زیادی از پدر بزرگ عموی عمهاش به او ارث رسیده بود.





گل مراد که از عکسش مشخص است مردی عاقل و فهمیده بود!!! برای اینکه این ثروت از بین نرود از قاضی شهر خواست تا او را راهنمایی بنماید! قاضی به او گفت: «گل مراد! تو باید صاحب فرزندی بشوی تا حافظ مال و ملال تو باشد...» گل مراد نیز که تا آن روز فرزندی نداشت شب به مسجد رفت و از خدا خواست تا به او فرزندی عطا بنماید.

این داستان تا ۳۲ سال ادامه داشت و هر روز گل مراد در مسجد به دعا مینشست اما خبری از فرزند نبود! در یکی از شبهایی که گل مراد بسیار دل شکسته و ناراحت بود فرشتهای بر او نازل گشت و به او گفت: «**گل مراد! ما دیگر از دست تو آرامش نداریم! برای همین آمدهام تا سخنی با تو بگویم**...» گل مراد که دید بلاخره دعاهایش جواب داده است با دقت گوش میداد که ناگهان فرشته گفت: «گل مراد جان! تو باید ابتدا ازدواج بنمایی!!!!!!!» در نتیجه گل مراد ازدواج کرد و نتیجه ازدواج گل مراد در کنار او ایستاده است و کل ثروتش نیز نماییده شد...

بله بچههای عزیز امیدوارم از این داستان متوجه شده باشید که **برای کسب نتایج عالی** در گوگل ابتدا باید ایندکس شوید. (یا شاید هم ازدواج کنید!) صفحاتی که ایندکس نشده باشند، چطور میتوانند در نتایج جستجو نمایش داده شوند؟ اگر گوگل صفحهای را ایندکس نکند هرگز هیچ رتبهای نخواهیم داشت!

بله! دقت کردید این سناریو به کجا رسید؟ **سئو سایت ما کاملا خاموش میشود** و هیچ راهی بجز بهبود ایندکس برای ما باقی نخواهد ماند! اهمیت این موضوع آنقدر بالاست که ما برای شما همین کتابی را که در حال مطالعه آن هستید به علاوه چندین مقاله کامل و مفید (همگی آنها را در فصل منابع مفید معرفی کردهایم) و یک دوره جامع همراه با مثالهای عملی را تهیه و منتشر کردهایم...

مسئله فقط مربوط به ایندکس نشدن سایت نیست؛ سایت ما ممکن است با سرعت و تعداد بالایی ایندکس شود؛ اما **مسئله مهم ما ایندکس صفحات مهم و ارزشمندمان هستند**! اگر بودجه محدود ما صرف بررسی و ایندکس صفحات بیارزش، کم اهمیت و بیکیفیت سایت ما شود، <mark>دیگر منابعی برای صفحات ارزشمند ما باقی نخواهد ماند</mark>. حالا وقت آن است تا با نحوه کار بودجه خزش آشنا شویم و آستینها را بالا بزنیم!







بخش ا میرزا ربات داروغه خراج چه صفحاتی را میگیرد؟

شاید فکر کنید بودجه خزش فقط مربوط به صفحههای سایتمان است اما اگر این تصور را دارید کاملا در اشتباه هستید! <mark>تمام آدرسها و درخواستهای سایت در بودجه</mark> خزش محاسبه میشوند</mark>؛ صفحاتی مانند فایلهای PDF، آدرسهای جایگزین (feed)، صفحات AMP، Hreflang، SSS، جاوا اسکریپتها و نسخه دسکتاپ و موبایل سایت از جمله این درخواستها هستند.

این آدرسها ممکن است با خزیدن و تجزیه و تحلیل صفحات یا از منابع دیگری مثل نقشـه سـایت، فیـدهای RSS، ارسـال آدرس در ابـزار گوگـل سـرچ کنسـول یـا اسـتفاده از API پیدا و خزیده شوند.

البته اصلاً نگران این مواردی که خواندید و احتمالاً گیج شدهاید نباشید، فقط قصد داشتیم شما بدانید فقط صفحات، در بودجه خزش محاسبه نمیشوند و موارد دیگری نیز هستند؛ در ادامه تمامی این موارد را تکبهتک بررسی و در ابزار گوگل سرچ کنسول آنها را تجزیه و تحلیل میکنیم.

اصلیترین و مهمترین مسئلهای که باید بدانید، قانون تعیین بودجه خزش است. **تعیین این بودجه حساب و کتاب خاص و دقیق خودش را دارد و گوگل روی میزان آن بسیار حساس است**. به هر حال <mark>خزیدن و درک محتوای هر صفحهای از وب هزینههای</mark> زمانی و مالی را به دنبال دارد.

در ادامه بهطور دقیق و مفید شما را با نحوه محاسبه بودجه خزش آشنا میکنیم. شما باید با بهینهسازی این موارد، به افزایش بودجه خزش هدفمند خود کمک کنید.

فش ۲ تعیین میزان بودجه خزش در مجلس شورای رباتی!

عوامل و فاکتورهای پیچیدهای در تعیین این بودجه تأثیرگذار هستند که همگی آنها کاملاً شفاف و مشخص نیستند؛ با این وجود چندین مورد کاملاً مشخص هستند و قصد داریم آنها را بررسی کنیم. این عوامل از تأثیرگذارترین و مهمترین شاخصها برای تعیین میزان بودجه خزش هر سایت هستند:

- ۱. بزرگی و تعداد صفحات سایت
 - ۲. قدرت پاسخگویی سرور
 - ۳. تازگی و کهنگی محتواها
 - ۴. سرعت بارگذاری صفحات
 - ۵. ساختار پیوندها
 - ۶. میزان سلامت سایت





هرکه سایتش بیش، خزش نیز بیشتر!

اگر سایت بزرگی داریم و تعداد صفحات آن از میزان مشخصی (مثلاً هزار صفحه) بیشتر باشد قاعدتاً **میزان بودجه خزش سایت نیز نسبت به سایتهای کوچکتر بیشتر است تا قدرت بررسی و ایندکس سایت افزایش یابد**؛ بنابراین اگر سایت کوچکی داریم نگرانی ما بابت میزان بودجه خزش نیز کمتر است؛ زیرا تمام صفحات سایتهای کوچک با همان میزان پایین بودجه نیز بالاخره روزی خزیده میشوند. (این به آن معنی نیست که اگر سایتمان کوچک است نگران بودجه خزش نباشیم ...)

مورچه چیه که کله پاچهاش چی باشه!

همانطور که در فصل قبل نیز به این نکته اشاره کردیم، قدرت سرور در تعیین میزان بودجه خزش تأثیر گذار است. رباتهای گوگل ابتدا حداکثر میزان درخواستهای خود را به سمت سرورهای ما روانه و میزان پاسخگویی سرور را نسبت به فشاری که وارد کردهاند بررسی میکنند. اگر به سرور ما فشاری وارد نشود و بتواند بهراحتی و درستی تمامی درخواستها را پاسخ دهد، یک سیگنال مثبت برای افزایش میزان بودجه خزش به گوگل ارسال میشود و همینطور بالعکس اگر سرور ما قدرت پاسخگویی به این میزان فشار وارده را نداشته باشد، قطعاً گوگل میزان بودجه خزش سایت را کم میکند تا خدایی ناکرده آسیبی به سرور و سایت وارد نکند.

نکتـه: البتـه مـا هـم میتوانیم میـزان فشـاری را کـه رباتهـای گوگـل بـه سـرورمان وارد میکنند محدود کنیم که در فصلهای آینده کاملاً به آن میپردازیم.

جوانیام بهاری بود و بگذشت... :(

در شرکت گوگل یک ضربالمثل هست که دیگر تبدیل به الگوریتم مهمی شده است! رباتهای گوگل هار روز بعاد از صارف صابحانه و هنگامی کاه قصاد دارناد باه ساراغ کارهایشان بروند، قبل از خداحافظی، این ضربالمثل را به یکدیگر یادآوری میکنند: «داداش! نو که اومد به بازار، کهنه میشه دل آزار! مراقب کدهات باش...»



سؤالهای مهم! مطالب سایت ما هرچند وقت یکبار بهروز میشوند؟ به نظر شما چند نفر به دنبال محتوای «معرفی بهترین گوشیهای هوشمند سال ۱۳۹۲» هستند؟ شاید در سال ۹۲ این محتوا بیش از چند میلیون بازدیدکننده را جذب کرده است، اما در حال حاضر دیگر ارزشی ندارد و روی سنگ قبرش این نوشته برق میزند: «تا محتوا بهروزه، کبکش خروس میخونه... گوگل چه مهربونه ... اما به وقت پیری، بییار و بینشونه ... باید تنها بمونه ... رتبه بقا نداره، الگوریتم حیا نداره گوگل وفا نداره ...»

شتاب صفر تا صد سایت: ۱.۴ ثانیه!

زمان ارزشمند است! چـه بـرای مـا انسـانها و چـه بـرای رباتـهـای بیاحسـاس ... مـا بـا افزایش سرعت بارگذاری و پاسخگویی سایتمان با یک تیر دو نشان میزنیم:

۱. در ابتدا تجربه بسیار لذتبخشی را برای کاربران بیحوصله این روزهای اینترنت رقم میزنیم.

۲. گوگل هم لذت میبرد و بودجه خزش ما را بیشتر میکند.

به نظر شما صفحهای با سرعت بارگذاری بسیار پایین، چه میزان زمان ارزشمند یک

ربات گوگـل را هـدر مىدهـد؟ اگـر صـفحات سـايت بـا سـرعت بـالايى بارگـذارى شـوند،

رباتهای گوگل زمان بیشتری را برای بررسی و خزش در صفحات، به دست میآورند<mark>.</mark>

رباتهای گوگل بسیار سرشان شلوغ است و مثلاً فقط ۵ دقیقه زمان را صرف بررسی سایت ما میکنند.

اگر سرعت بارگذاری هر یک از صفحات سایت به طور میانگین ۱ دقیقه باشد، ربات گوگـل فقـط فرصـت بررسـی ۵ صـفحه از سـایت را دارد. حـال اگـر مـا سـرعت بارگـذاری صـفحات سـایت را بهینـه کنـیم و آن را بـه ۳۰ ثانیـه برسـانیم، ربـات گوگـل در زمـان ۵ دقیقه، میتواند ۱۰ صفحه از سایت را بررسی کند.

یعنی **با ۲ برابر کردن سرعت سایت، صفحات سایت ۲ برابر بیشتر خزیده میشود**؛ در نتیجه بودجه خزش سایت نیز بیشتر میشود.



رباتی که دیوانه شده است!

اگر مثل من یک بازیکن حرفهای بازی مافیا هستید،حتما میدانید کسی که در این بازی، صحبتهای کم ارزشی وارد بازی کند و حرف حسابی برای گفتن نداشته باشد، قطعاً یا مافیاست یا حرفهای او برای ما اهمیتی ندارد؛ در نتیجه ما به افرادی که حقیقتهایی درست و منطقی را وارد بازی میکنند توجه بیشتری میکنیم. <mark>اگر</mark> لینکهای سایتمان به صفحات بیارزش، هرزنامه، محتواهای تکراری و از این دست صفحهها داده شود، قطعاً ما از نظر گوگل مافیای فضای اینترنت هستیم و دیگر لینکهای موجود در صفحات سایتمان برای او جذاب نیست و توجه کمتری به آنها میکند! و این یعنی کاهش بودجه خزش ...

همچنین **اگر ساختار لینکسازی سایت، غیر منطقی باشد** و باعث گمراهی و گیج شدن رباتهای گوگل شود، حوصله رباتها سر میرود و خزش این سایت برای آنها تبدیل به کابوس شبهایشان میشود و خواب و خوراک را از آنها میگیرد. فرض کنید ما به صفحه «خرید بستنی زعفرانی» لینک دادهایم، بعد از مدتی دیگر قرار نیست این محصول را بفروشیم، بنابراین این صفحه را با استفاده از ریدایرکت ۳۰۱ (در ادامه با این مفهوم آشنا میشویم) به صفحه «طرز تهیه بستنی زعفرانی» منتقل میکنیم. اما این صفحه رتبه خوبی نمیگیرد و برای ما نیز دیگر ارزشی ندارد؛ بنابراین تصمیم میگیریم آن را حذف کنیم. حالا چه اتفاقی افتاد؟ ربات گوگل پس از چندین بار جابهجایی بین صفحات، به یک بنبست رسید! این یعنی وقوع یک فاجعه و



به پرتگاه غم رسیده گامهای من، تاریکم! فردا سراغ من بیا ...

این که دیگر کاملاً مشخص است! در بهترین حالت، من به ماشین خرابم که چراغ چک آن با پررنگترین رنگ زرد دنیا روشن است، توکل کنم و برای خرید چیپس و ماست موسیر و یک بطری آب آلبالو تا سر کوچه بروم و برگردم (شاید هم برنگردم و به دلایلی که به خودم مربوط است جواز برگشتم فقط فیش حقوقی باشد!) قطعاً با این ماشین از شیراز به تهران نمیروم و حتماً قبل از سفر، تمامی ایرادهایش را بررسی میکنم و از سلامت کامل آن مطمئن میشوم.

وجـود خطاهـای متنـوع و محـدودیتهایی کـه بـه اشـتباه در سـایت وجـود دارد، روی سـلامت سـایت تأثیرگـذار اسـت و آن را مـریض میکنـد</mark>. چـه کسـی از مـریض شـدن خوشش میآید که رباتهای گوگل خوششان بیاید؟

آیا شعر شاعر محبوب و مشهور رباتها «ویلیام قطعهای از ربات را تکان میدهد» (William Shake spare a bot) که در ابتدای این فصل به آن اشاره کردیم را به یاد دارید؟

سایتی کـه خطاهـای اساسـی و مهمـی در آن وجـود دارد، بودجـه خـزش کمتـری را نیـز دریافت میکنـد و همـین دلیـل کـافی است تـا کمی هـم بـه فکـر پشت پـرده سـایتمان باشیم.

بهترین دکتر برای سایت ما، گوگل سرچ کنسول (Google Search Console) است. تمامی خطاها و ایرادهای سایت را در این ابزار میتوانیم مشاهده و بهترین نسخههای درمانی را برای اصلاح آنها دریافت کنیم؛ نگران نباشید در فصلهای آینده شما را بیشتر با این ابزار آشنا میکنیم.



بخش المان بندى خزيدن: واقعا دود از كنده بلند مى شود؟!

به نظر شما محتوای «قیمت لحظهای دلار» با محتوای «تاریخچه جنگ جهانی دوم» چه تفاوت آشکاری دارد؟ مگر تاریخچه جنگ جهانی دوم تغییری هم میکند؟ اگر تغییری هم کند چند بار این اتفاق خواهد افتاد؟ بنابراین گوگل هم از این موضوع اطلاع کافی دارد و میداند **قرار نیست تمام محتواها تازه باشند و بهروز شوند**؛ به تاریخ انتشار دو محتوای زیر دقت کنید:

> JOAXr ‹ https://www.aparat.com › دقیقا_چرا_laquo&:جن... دقیقا چرا «جنگ جهانی دوم» اتفاق افتاد؟ - آپارات



بدون دونستن **جنگ جهانی** اول و اتفاقات اون **جنگ**، ما نمیتونی به هیچ عنوان ... ما در **جنگ جهانی** اول گفتیم حتی دستور یک حمله به فرانسه... آیارات ۱۰ تیر ۱۴۹۱

inews < https://www.hamshahrionline.ir ۲ آشنایی-با-جن... * آشنایی با جنگ جهانی اول (۱۹۱۴-۱۹۱۸) - همشهری آنلاین ۱۸ مرداد ۱۳۸۷ – همشهریآنلاین: اتریش و مجارستان در ۲۸ ژوئیه ۱۹۱۴به صربستان اعلان جنگ کردند و این جنگ باعث به راه افتادن جنگی عالمگیر شد.

محتوای اول با تاریخ ۱ تیر ۱۴۰۱ و محتوای دوم با تاریخ ۱۸ مرداد ۱۳۸۷ منتشر شده است؛ این دو محتوا ۱۴ سال اختلاف زمانی دارند؛ اما بهخوبی و خوشی در صفحه اول گوگل باهم رقابت میکنند. گوگل اگر صد بار دیگر هم این دو محتوا را بررسی کند تغییری در آن پیدا نمیکند؛ بنابراین گوگل برای به هدر نرفتن منابع خود، مفهومی را به نام «زمانبندی خزش» معرفی میکند.

حال آن طرف داستان را نیز باهم بررسی کنیم؛ محتوای منتشر شده دیروز درباره قیمت لحظهای دلار، امروز ارزشی دارد که محتوای ۱۴ سال پیش داشته باشد؟؟ قیمت دلار ثانیه به ثانیه در حال نوسان است و گوگل باید این صفحه را لحظه به لحظه بررسی کند تا قادر به ارائه بهترین پاسخ برای کاربرانش باشد.

زمانبندی خزش برای تعیین ارزش خزیدن مجدد صفحات است؛ بهصورت سادهتر مشخص میکند چـه صفحهای و در چـه مـدت زمـانی ارزش دوبـاره بررسـی و اینـدکس شدن را دارد.



نکته بسیار مهم: <mark>بهطورکلی محتواهای تازه، اولویت بیشتری نسبت به صفحات قدیمی</mark> <mark>دارند</mark>. در همان مثال تاریخچـه جنـگ جهـانی دوم، محتـوایی کـه بسـیار قـدیمی *ب*ـود در لینک هفتم قرار دارد و محتواهای جدیدتر نسبت به آن رتبه بالاتری دارند.

زمانبندی خزش بر اساس چه معیارهایی تعیین میشود؟

اگر گوگل صفحهای را بررسی کند و متوجه شود محتوای آن تغییری نمیکند، کمتر به آن سر میزند. برای مثال اگر گوگل یک صفحه را بخزد و در روز بعد هیچ تغییری در آن نبیند، ممکن است ۳ روز بعد دوباره به سراغ آن بیاید و اگر باز هم تغییری نبیند ۱۰ روز دیگر دوباره آن را بررسی کند و همینطور بازههای زمانی بررسی را به ۳۰ – ۵۰ – ۱۰۰ و بیشتر به تأخیر بیندازد.



۴ مورد از مهمترین معیارهای بررسی دوباره صفحات شامل این موارد است:

- ۱. نوع صفحه
- ۲. نرخ بهروزرسانی و تازگی
 - ۳. محبوبیت صفحه
 - ۴. اعتبار دریافتی

اونی که همش عوض میشه، میخواد بگه بخز منو، روش نمیشه!

یک سایت از چندین نوع صفحه تشکیل شده است؛ صفحاتی مانند:

- ارتباط با ما
- شرایط و ضوابط
 - درباره ما
- آرشیو و صفحه تکی مقالات
 - دستەبندى محصولات
 - لندینگها و ...

صفحه دستهبندی محصول «گوشی موبایل سامسونگ» و صفحه «روند ثبت سفارش» را در سایت دیجیکالا در نظر بگیرید، **کدام صفحه تغییر بیشتری میکند و ارزش بیشتری برای خزیدن دوباره را دارد؟** بنابراین نوع صفحه نیز در زمانبندی خزش تأثیرگذار است.

صفحات تازه، خوشمزهتر هستند!

همانطور که نوشتیم، صفحههایی با نرخ بهروزرسانی بالاتر، پتانسیل بیشتری را برای خزیده شدن نیاز دارند.



دل منو بردی!

صفحات محبوب نیاز به بررسیهای بیشتری دارند، این محبوبیت بر اساس چند فاکتور تعیین میشود:

- میزان تعامل مخاطبین
 - نرخ بازدید محتوا
 - رتبه کنونی صفحه
- میزان محبوبیت کلی سایت

پارتیبازی در سایت

هر چقدر تعداد بکلینکهای (لینکهای ورودی به سایت) یک صفحه از سایت بیشتر باشد، اعتبار بیشتری نیز دارد. این صفحه نه لزوماً اما احتمالاً نسبت به سایر صفحات بیشتر خزیده میشود

بهعنوان نکته پایانی: سعی کنیم اطلاعات مفیدتری را برای غنیسازی محتواهای خود، در طول زمان به آنها اضافه کنیم. <mark>توجه داشته باشیم این اطلاعات باید تمرکز موضوع</mark> محتوایی ما را حفظ کند و برای کاربر نیز ارزشمند باشد<mark>.</mark>

تبریک میگویم شما تا همینجای کتاب، کاملاً با مفهوم بودجه خزش آشنا شدید و با این تلاشهایی که من برای شیرینکردن محتوا انجام دادهام مطمئنم خسته نیستید و کاملاً آمادهاید تا با یک مفهوم دیگر آشنا شوید:

+ آمادەايد بچەھا؟

- بللله ناخدا :)





شناخت مفاهیم مرتبط در حل چالشها و مسائل سئو بسیار یاری رسان و ارزشمند است. در این فصل شما با ۱۱ مفهوم بسیار مهم به صورت خلاصه، مفید و بامزه آشنا خواهید شد. اگر این مفاهیم را میشناسید، پیشنهاد میدهیم از این فصل رد نشوید و حتما چند دقیقه را برای مطالعه دوباره آنها زمان بگذارید.

- ۱. نقشه سایت
- ۲. فایل دستور العمل رباتها
 - ۳. بازرس آدرس
 - ۴. تگ کنونیکال
 - ۵. ریدایرکت ۳۰۱
 - ۶. محتوای تکراری
 - ۷. لینک شکسته
 - ۸. لینک داخلی
 - ۹. صفحات یتیم
 - ۰۱. تگ نو ایندکس
 - ۱۱. کنیبالیزیشن





بخش ا نقشه سایت: آدرس سر راست است!

همانطور که سفر به دنیایی ناشناخته، بدون همراه داشتن نقشه راهنما دشوار است، درک معماری و یافتن صفحات مهم و ارزشمند سایت نیز ممکن است برای گوگل سخت و گیچکننده باشد. در این شرایط، نقشه سایت میتواند حلال این چالشها، میان ما و گوگل باشد! نقشه سایت یک فایل XML است که تمام محتواهای مهم سایت را فهرست میکند و رباتهای گوگل با استفاده از آن متوجه میشوند که چه محتواهایی در سایت وجود دارند و چگونه میتوانند به آنها دسترسی پیدا کنند. نقشه سایت یک نمای کلی از تمام محتواهای موجود و ارزشمند را به طور یکجا و همزمان در اختیار رباتهای گوگل قرار میدهد و به همین دلیل، اهمیت زیادی برای ما و گوگل دارد! در نقشه سایت علاوه بر ارائه صفحات ارزشمند و مهم، امکان افزودن

دارد.

نحوه دسترسی به نقشه سایت

فایل xml نقشه سایت معمولاً بلافاصله بعد از نام دامنه قرار دارد. مثلاً به این آدرس توجه کنید: https://www.websima.academy/**sitemap.xml**

کافی است بهجای عبارت websima.academy نام دامنه سایت و بهجای عبارت sitemap نام نقشه سایت را قرار دهید. **مدیر سایت، نام نقشه سایت را انتخاب میکند و ممکـن اسـت هـر چیـزی باشـد** امـا بـه طـور کلـی نقشـه سـایت بـا نـام سـایت مـپ (sitemap) ذخیره میشود؛ پس اگر نام آن را بدانیم با جستجوکردن، نقشه سایتمان را پیدا میکنیم.

در فصلهای بعدی، استانداردها و نحوه بارگذاری نقشه سایت روی سایت و معرفی آن به گوگل را بررسی میکنیم.

نقشه سایت چیست؟ ساخت سایت مپ و معرفی به گوگل 🧩

بخش ٢ دستور العمل رباتها: فضول نباش ربات جان!

فایل دستورالعمل رباتها یک فرشته نجات برای سایت است! این فایل دستوراتی را به رباتهای موتور جستجو میدهد و سایت ما را از فاجعه نابود شدن بودجه خزش نجات میدهد.

یک سایت فروشگاهی را تصور کنید که قابلیت ایجاد صدها صفحه را با فیلترهای مختلف دارد. بـهعنوان مثـال فیلتـر رنگهـای صفحه «کفـش نایـک» را بـا وجـود ۱۰۰ محصول در نظر بگیرید؛ اگر در این صفحه فیلتر رنگ آبی را فعال کنیم، صفحهای با آدرس جدیدی ایجاد میشود که ۸۰ محصول با رنگ دیگر را حذف کرده و ۲۰ عدد کفش نایک آبیرنگ را نشان میدهد. (دقت کنید که محتوای صفحه تغییری نکرد و فقط محصـولات دیگـر حـذف شـدند) ایـن صفحه کـار مخـاطبین را بـرای بررسـی و مقایسـه کفشها بسیار راحت میکند؛ اما یک مشکل بزرگ در پشت پـرده ایـن صفحه وجود دارد:

با انتخاب فیلتر، آدرس صفحه تغییر میکند؛ اما محتوای جدید و ارزشمندی داده نمیشود. صفحات فیلتر شده باعث ایجاد محتواهای تکراری میشوند و همین کافی است تا رباتهای گوگل در تشخیص صفحات مهم و ارزشمند سایت گیج شوند و فکر کنند محتواهای ما تکراری هستند و ارزشی ندارند.



ما باید قوانینی تنظیم کنیم تا **رباتهای گوگل به این نوع صفحات دسترسی نداشته باشند** و همچنین اگر محتوای صفحه را بررسی کردند آن را ایندکس نکنند. برای این کار میتوانیم از فایل دستورالعمل رباتها (robots.txt) استفاده کنیم.

فایل (robots.txt) مجموعهای از دستورالعملهای اختیاری برای موتورهای جستجو است. این فایل به خزندههای گوگل میگوید که در بعضی از صفحات یا بخشهای خاصی از سایت نخزند. (یا اگر خزیدند، آنها را ایندکس نکنند.)

خبر خوب این است که اصلیترین موتور جستجو (گوگل) درخواستهای Robots.txt را میشناسد و به آنها احترام میگذارد. رباتهای گوگل قبل از بررسی سایت، ابتدا به دنبال فایل (robots.txt) میگردند و **اگر آن را پیدا کنند دستورات آن را میخوانند** و سپس کار خود را شروع میکنند؛ بنابراین خیالتان راحت باشد که با استفاده از این فایل، دیگر مشکلی پیش نخواهد آمد.

نمونه آدرس robots.txt: این فایل باید کجا و چگونه قرار بگیرد؟

قبل از شروع مفهوم بعدی، سه نکته مهم را از قلم نیندازیم!

فایل robots.txt همیشه باید با حروف کوچک، فرمت txt و بلافاصله بعد از آدرس سـایتمان قـرار بگیـرد. (websima.academy/robots.txt) اگـر ایـن اسـتانداردها را رعایت نکنیم، رباتهای گوگل آن را پیدا نمیکنند.

فایل Robots.txt چیست؟ تابلو ورود ممنوع برای رباتهای گوگل 💥



بخش ۳ URL Inspection: بازپرس دادگاه سرچ کنسول

خدا میداند که تشخیص مسائل فنی در وبسایتها، چقدر زمانبر است و این کار آنقدر اهمیت دارد که نمیشود بیخیال آن شد؛ خوشبختانه مجبور نیستیم برای این کار یکعالمه وقت بگذاریم. با استفاده از ابزار بازرسی آدرس «URL Inspection» در گوگـل سـرچ کنسـول میتـوانیم بسـیار آسـان و سـریع صـفحات سـایتمان را بررسـی و اشکالات فنی آن را رفع کنیم.

کلمـه «Inspection» در زبـان انگلیسی بـه معنـای «بازرسی» است. (واقعـاً هـم یـک «بازرس» بـه تمام معناست!) ایـن ابزار آ**درسهای سایت را بررسی میکند و اطلاعات صـفحات را بـه صـورت لیسـت در اختیارمـان قـرار میدهـد**. (یـک بازرسـی عـالی و بیعیبونقص که کارمان را حسابی راحت میکند) مثلاً متوجه میشویم آیا صفحات سایتمان به درستی ایندکس شدهاند و برای کاربران نمایش داده میشوند یا نه؟ اگر یکی از صفحات ما ایندکس نشده باشد، با بررسی اطلاعات این ابزار متوجه خواهیم شد اشکال کار کجاست و چطور باید آن را رفع کنیم.

اگر تا همینجا پیش خودتان فکر کردهاید که «بهبه، چه ابزار بینظیری! برم و کلی باهـاش کیـف کـنم و لـذتش رو ببـرم …» بایـد بگـوییم کـه تـازه کجـای کـار را دیدهایـد! بازرس ما چند مزیت بسیار کارآمد دیگری را نیز در چنته دارد:

- فایل دستورالعملهای رباتها (robots.txt) را که یادتان نرفته از بازرس، قوانین
 این فایل را نیز بررسی میکند تا اگر خدای ناکرده بهدرستی تنظیم نشده باشد،
 این موضوع را فوراً به ما اطلاع دهد.
- بازرس مثل یک کاراگاه لیست صفحههایی را که رباتهای گوگل آنها را بررسی
 کردهاند و لینک این صفحه در آنها بوده است را ارائه میدهد.
- فورا نقشه سایت را بررسی میکند و به ما از وجود یا عدم وجود صفحه در آن
 اطلاع میدهد
- اطلاعات مربوط نشانه گذاریهای صفحه مانند: سؤالات متداول، ویدئو، محصول،
 محتوا و ... را بررسی میکند و...

URL Inspection؛ سریعترین راه ایندکس دستی صفحه در گوگل

^{بخش ۴} تگ کنونیکال: اصلش رو بدم؟ فیکش رو بدم؟

<"/link rel="canonical" href="https://www.websima.acadmy/" مثال بالا یک تگ کنونیکال است. امیدوارم یاد گرفته باشید؛ **تا مباحث و مفاهیم بعدی شما را به خدای مهربان میسپارم**. نویسنده: «چرا هنوز داری میخونی؟ همین بود دیگه! تمـوم شـد ... **مگـه نمـیگم نخـون تمـوم شـد!!!** عـه عـه عـه! بـاز هـم داری میخونی؟؟؟؟ نه خوشم اومد، خیلی دوست داری یاد بگیری؛ **باشه خوشگل من!** ادامه میدیم ...»

تگ کنونیکال کاربردی بسیار اساسی و مهمی دارد و نقش آن در سایت حیاتی است؛ این تگ <mark>آدرس بهترین نسخه از محتوا را به گوگل ارائه میدهد</mark> و با همین کار ساده از ایندکس شدن صفحاتی با آدرسهای متفاوت و محتوای یکسان جلوگیری میکند.

حالا چرا این تگ اهمیت دارد؟

اگر ما صفحههایی را ببینیم که ظاهر و محتوای آنها تفاوتی نداشته باشند، فرقی بین آنها قائل نمیشویم. مثلاً صفحه اصلی سایت ما همیشه یک چیز را نشان میدهد حتی اگر با آدرسهای مختلفی باز شود. اما این ماجرا برای رباتهای گوگل فرق میکند؛ رباتهای گوگل هر آدرسی را بهعنوان یک صفحه جداگانه میشناسند، حتی اگر ظاهرشان و تمام محتوایشان یکی باشد.

✻

شاید الان این سؤال برایتان پیش آمده که «مگر میشود یک صفحه با چندین آدرس متفاوت در دسترس باشد؟» برای همین ما چند آدرس صفحه اصلی آکادمی را برای شما مثال میزنیم تا این موضوع را بهتر درک کنید:

- https://websima.academy/ •
- https://websima.academy/amp •
- https://websima.academy/feed •
- https://websima.academy/print •

اگر دقت کرده باشید وجود صفحاتی مانند feed ،AMP و... باعث ایجاد یک آدرس جدید با محتوای تکراری میشوند؛ بنابراین چندین آدرس مختلف برای صفحه اصلی سایت ما وجود دارد که هرکدام از آنها در نظر گوگل صفحهای جداگانه هستند.

این اتفاق مخصوصا **در سایتهای فروشگاهی به دلیل وجود گزینههای فیلتر، بسیار رایج است**. مثلا در صفحه خرید ساعت مچی با انتخاب فیلتر جدیدترین محصولات یا مرتبسازی بـر اسـاس گرانتـرین، یـک پـارامتر بـه آدرس صفحه اضـافه میشـود، امـا محتوای صفحه تغییر نمیکند.

در حال حاضر چندین آدرس مختلف با محتوای یکسان وجود دارند که باعث میشوند، بخشی از سایت ما از نظر گوگل تکراری باشد، بودجه خزش ما را هدر دهند و سایت ما نابود شود. راه حل چیست؟ درست حدس زدید؛ باید از تـگ کنونیکال در سایر آدرسهای تکراری استفاده کنیم. <mark>ربات گوگل با دیدن این تگ متوجه آدرس اصلیترین</mark> و بهترین نسخه محتوا خواهد شد و فقط آن را ایندکس میکند. به همین راحتی ... با این حساب باید خدا را شکر کنیم که تگ کنونیکال از اتفاقات بسیار بدی جلوگیری میکند و ما را از کلی دردسر نجات میدهد!

تگ کنونیکال؛ راه خروج از هزارتوی محتوای تکراری



9

×

بخش ه ریدایرکت ۱۰۳: این به سر کوچه انتقال یافت!

مـن (نویسـنده) در طـول ۳ سـال گذشـته، ۵ بـار خانـهام را عـوض کـردهام و در آدرس جدیدی زندگی میکنم؛ اما هر سری، ثبت آدرس جدیدم را فراموش میکنم و به همین دلیـل سـفارشها، هدیـهها و مرسـولههای پسـتیام معمـولاً بـه آدرس قبلـی ارسـال میشود و این ماجرا باعث ایجاد چالشهای عجیبی شده است.

این بهترین اتفاق بدی بود که کاملاً مناسب اینجاست! آقای پشمک را که خاطرتان هست؟ فرض کنید آقای پشمک روی لینکی کلیک میکند که آدرس صفحهاش تغییر کــــرده اســــت (مـــــثلاً آدرس websima.academy/workshop را بـــه ناکام؛ خدا رحمتش کند!) روبرو میشود. در این صورت، قطعاً حسابی توی ذوقش میخورد و به سراغ سایت دیگری میرود؛ اما ما دلمان نمیخواهد آقای پشمک را ناامید کنیم، **ما آقای پشمک را دوست میداریم**، آقای پشمک جزوی از برندینگ ماست، ما نمیخواهیم آقای پشمک سراغ سایتهای رقیب برود!!!

پـس از ريـدايركت ۳۰۱ اسـتفاده مـىكنيم و اصـطلاحاً صـفحه «workshop/» را روى صفحه «course/» ريـدايركت ۳۰۱ مىكنيم تـا آقـاى پشـمك بـه طـور خودكـار بـه آدرس جديد منتقل شود. به همين راحتى!

کاربران ما اصلاً روحشان هم خبر ندارد که آدرس ما عوض شده است و اصلاً نمیدانند که آدرس جدید ما چیست! بنابراین اگر آدرس یکی از صفحات ما تغییر کند، باید آدرس قبلی را به آدرس جدید منتقـل کنیم تا بازدیدکننـدگان ما به درستی وارد آدرس جدید شوند.

ریدایرکت ۳۰۱ چیست؟ شمارش معکوس از دست دادن رتبه!

9

×

بخش ۶ محتوای تکراری: تکرار قسمت ۲ جومونگ ساعت ۱۴

یکی از بزرگترین افسانههای ترسناک در مباحث سئو «جریمه سایت به علت وجود محتوای تکراری» است! (باور کنید که این عبارت برای صاحبان سایت، از هـر فیلم ترسناکی وحشتناکتر است!) اما میخواهید یک راز سئو را بدانید؟ «**چیزی بهعنوان جریمه محتوای تکراری وجود ندارد!**»

اما منظورمان این نیست که داشتن محتوای تکراری ضرری برای سایت ما ندارد. داستان از این قرار است که وقتی گوگل با محتواهای یکسانی در یک یا چند سایت مواجه میشود، باید تصمیم بگیرد که به کدام محتوا رتبه بالاتری بدهد. از بخت بد، همیشه این رتبهبندی در شرایطی عادلانه انجام نمیشود. در نتیجه ممکن است محتوای مفیدی که بهسختی روی آن کارکردهاید، مثل یک روح نامرئی شود و اصلاً کسی آن را نبیند!

خبر بـد ایـن اسـت کـه بسـیاری اوقـات مـا کنترلـی روی تکـراری شـدن محتوای سـایتمان نداریم. مثلاً ممکن است فرد دیگری محتوای ما را در سایت خـودش کپـی کنـد و همـین موضوع باعث شود که محتوای سایت ما تکراری بهحساب بیاید.

دقت داشته باشید که <mark>محتوای تکراری فقط به صورت متن نیست و میتواند صفحات</mark> جستجوی سایت، فیلترها، محتواهای مشترک در چند دستهبندی و ... نیز باشد<mark>.</mark>

محتوای کپی چیست؟ چرا همه درباره آن اشتباه میکنند؟



✻

س 🕐 لینک شکسته: بودن یا نبودن! مسئله این است...

بگذارید در یککلام این عبارت را تعریف کنیم: <mark>لینک شکسته (Broken Link) لینکی</mark> است که کار نمیکند یا خراب است وبسایت کار نکرد و شما با وارد شدن در آن، با یک صفحه ارور بزرگ مواجه شدید، اصطلاحاً میگوییم آن لینک، شکسته است. برخی از عوامل ایجاد لینک شکسته این موارد هستند:

- وبسایت مورد نظر حذف شده است یا آدرس آن تغییر کرده باشد.
- صاحب سایت، صفحه مد نظر را خصوصی کرده باشد و کاربران دیگر امکان
 دسترسی به آن را نداشته باشند.
 - در آدرس اشتباه تایپی وجود داشته باشد
 - سرور سایت مد نظر از کار افتاده باشد

سایت Raven (فعال در حوزه سئو) با تحقیقی که در سال ۲۰۱۳ تا ۲۰۱۵ انجام داد، مشخص کرد که **خطر وجود لینکهای شکسته بیخ گوش همه ماست!** این تحقیق نشـان داد کـه در هـر سـایتی حـداقل حـدود ۲۳ لینـک شکسـته وجـود دارد، درحالیکـه مدیران سایت حتی روحشان هم از این موضوع خبر ندارد!

گوگل میگوید که وجود لینکهای شکسته در وبسایت ضرری به سئوی ما نمیرساند. ما این حرف را رد نمیکنیم، اما عقیده داریم که این لینکها با این که شاید دشمن ما نباشند، اما دوستمان هم نیستند. (**البته ما تعداد کمی لینک شکسته را در نظر داریم**) وجود این لینکها باعث میشود که کاربران، تجربه کاربری خوبی از سایت ما نداشته باشند و در نتیجه بازدیدکنندگان زیادی را از دست بدهیم. پس چه بهتر که همین الان دستبهکار شویم و از شر این مهمانان ناخوانده سایتمان خلاص شویم.



وبسایتها و ابزارهایی هستند که سایت ما را جستجو میکنند و لینکهای شکسته را تشخیص میدهند؛ برای مثال ابزار رنک مث (Rank Math) یکی از آن ابزارهای قابل اعتماد است؛ این ابزار با آرامش کامل مانند یک جاروبرقی، تمام لینکهای شکسته در سایت ما را جارو میکند و یک سایت تروتمیز و خالی از لینک شکسته تحویلمان میدهد.

البته اگر حرف من را قبول دارید (**که مطمئنم دارید!**) جعبه ابزارهای بسیار پیشرفتهتری هستند که کار با آنها شما را کاملاً حرفهای میکند و دیگر نیازی به چندین ابزار مختلف ندارید. مثلاً استفاده از ابزار اسکریمینگ فراگ بینظیر است و هر متخصص سئویی به آن نیاز دارد. (آموزش کامل استفاده از این ابزار را میتوانید در دوره سئو تکنیکال مشاهده کنید و با یک تیر هم سئو تکنیکال را یاد بگیرید و به صورت ویدئویی با مفهوم بودجه خزش آشنا شوید و هم با یادگیری این ابزار از شر تمام ارورهای سایت خلاص شوید.)

Broken Linkچیست؟ چه تاثیری بر سئو سایت خواهد داشت؟

پیشگیری همیشه بهتر از درمان است!

✻

9

یک راه عـالی بـرای پیشـگیری از ایجـاد لینکهـای شکسـته حـذف نکـردن هـیچ یـک از صفحات سایتمان است</mark>. اگر صفحاتی قـدیمی در سایتمان داریم که محتوایشان دیگر به دردمان نمیخورد، بهتر است جای آنها را با یک محتوای جدید و کارآمد عوض یا آن را به آدرس یکی از صفحات مرتبط (در صورت وجود) منتقـل کنیم. بـه ایـن صورت دیگر نیازی به حذف صفحات نخواهـد بـود و در نتیجـه از خطر ایجـاد لینکهـای شکسته هم دور میمانیم.



حالا باتوجهبه مفاهیمی که تابهحال یاد گرفتهاید میتوانید حدس بزنید که برای انتقال صفحات باید از چه ابزاری استفاده کنیم تا لینکهای سایت به لینک شکسته تبدیل نشوند؟ آفرین! **تگ کنونیکال!** کاملاً اشتباه حدس زدید!!! یک راه عالی برای پیشگیری از بهوجودآمدن لینکهای شکسته استفاده از ریدایرکت ۳۰۱ است ...

بخش ۸ لینک داخلی: پیدا کن مرا، شاید آشنا باشم با تو

اگر بخواهید مـن را در محـل کـارم پیـدا کنیـد، بـرای رسـیدن بـه مـن یکـی از ایـن دو راه امکانپذیر است:

- آدرس دقیق من را دارید و مستقیماً به سراغم خواهید آمد
- از همکار من میپرسید «ابوالفضل پارسائی را کجا میتوانم پیدا کنم؟»

فرض کنید شما آدرسی از من ندارید و من هم همکاری نداشته باشم، راه دیگری دارید؟ **قطعاً دیگر به من دسترسی ندارید**. لینکهای داخلی دقیقاً مانند همکاران من عمل میکنند. هرچقدر تعداد نفرات بیشتری من را بشناسند، علاوه بر اینکه شما بهتر و راحتتر مرا پیدا میکنید و به من دسترسی دارید، **اعتبار من نیز بیشتر میشود**. کاربران با استفاده از لینکهای داخلی میتوانند صفحات مدنظر خودشان را پیدا کنند و در سایت ما حرکت کنند، <mark>گوگل نیز با استفاده از لینکهای داخلی، ارزش و اعتبار</mark> صفحات ما را تخمین میزند و صفحات جدید را نیز پیدا و ایندکس میکند.

تمام لینکهایی که در منوی اصلی سایت، سایدبار، محتوای صفحه، بنرهای گرافیکی، فوتر و سایر بخشهای صفحه قرار میگیرند و مقصد آنها درون دامنه سایت است، در ارتباط بین صفحات تأثیرگذار هستند و از انواع لینکسازی داخلی محسوب میشوند. اگر بـرای صفحهای لینـک داخلـی نداشـته باشـد، تنهـا راه رسـیدن بـه آن داشـتن آدرس کامل آن است؛ **اما کاربران ما و گوگل این آدرس را ندارند**.

بنابراین بدون وجود یک استراتژی منطقی و منظم برای لینکسازی داخلی، سایت ما شبیه به یک هزار توی گمراه کننده است که نه برای گوگل و نه برای مخاطبین کارایی لازم را ندارد. لینکسازی داخلی یک مبحث بسیار مهم و گسترده است و آنقدر اهمیت داشت که ما یک دوره جامع و مفید برای آن تهیه و منتشر کردهایم.

لینکسازی داخلی چیست؟ رسیدن به رتبه یک یا سقوط به آخرین صفحه گوگل؟



9

آموزش حرفهای لینکسازی: کسب اعتبار از طریق لینکسازی و برندینگ

مش ۹ صفحات یتیم: هنوزم، لینکهای تو، برام مثل عمر دوباره است

متخصص سئویی پس از چندین سال پیر و کهـن سال شـد و در بسـتر مـرگ افتـاد، پسرهایش را صدا زد. هر هفت پسر آمدند و بر بالین پدر نشستند. پیرمرد چوبهایی را که از قبل آماده کـرده بـود از زیـر بـالش در آورد و بـه هـر کدامشـان یـک تکـه چـوب داد تـا بشکنند. **پسرها به راحتی آب خوردن چوبها را شکستند**.

پیرمرد لبخندي زد و گفت: «هنگامی که صفحهای، از سایت شما جدا باشد آن موقع دیگر هرگز ایندکس نخواهد شد و به راحتی شکست میخورد. حالا صبر کنید!» بعـد از زیـر بـالش مقـدار زیـادی چـوب بیـرون آورد و ایـن بـار بـه هـر کدامشـان یـک دسـته چـوب هفتـتایی داد و گفت: «اگر میتوانید حالا چوبها را بشکنید!»

میدانید چه شد؟ آفرین! این بار دیگر کاملاً درست حدس زدید!!! این دفعه هم پسرها مثل آب خوردن چوبها را شکستند؛ هر هفت پسر یتیم شدند و پدر هرگز نتوانست به پسرهایش اصول سئو را یاد بدهد...



اگر صفحهای از سایت ما جـدا افتـاده باشـد و از هـیچ صفحهای لینکی دریافـت نکنـد، اصطلاحاً به آن صفحه یتیم میگوییم.

این صفحه هرچقدر هم که خوب باشد به احتمال خیلی زیاد ایندکس نخواهد شد و ارزشی را برای سایت ما ایجاد نخواهد کرد. به نظر شما برای ایندکس شدن این صفحه از چه ابزاری میتوانیم کمک بگیریم؟

- بازرس آدرسها (URL Inspection)
 - نقشه سایت (Site Map)
- افزودن لینکهای داخلی (Internal Links)

البته **گاهی هم بهوجودآمدن این صفحات عمدی است!** مثلاً صفحات لندینگی که صرفا برای تبلیغات آنها را طراحی کردهایم، یکی از انواع صفحات یتیم محسوب میشوند؛ زیـرا فقـط تبلیغـات هسـتند و لینکسـازی داخلـی بـرای بعضـی از ایـن صفحات منطقـی نیست.

^{بخش ۱۰} تگ نو ایندکس: تو رو سننه قربونت برم!

کمی به عقب برگردیم: وقتی عبارتی را در گوگل جستجو میکنیم، **نتایجی به ما نشان** داده میشوند که به هیچ عنوان تصادفی انتخاب نشدهاند؛ انتخاب محتواهایی که در صفحه نتایج به ما نمایش داده میشوند به ۲ عامل مهم بستگی دارد:

- «یافتن»
- «رتبەبندى»



رباتهای خستگیناپذیر گوگل، بیوقفه محتواهای جدید را بررسی و ایندکس میکنند و با توجه به الگوریتمهای گوگل تصمیم میگیرند که هر محتوا در کجا قرار بگیرد و چه رتبهای داشته باشد.

حال اگر ما دوست نداشته باشیم محتواهایمان ایندکس شود و در نتایج جستجو بیایند، باید چه کاری انجام دهیم؟ یا شاید هم محتوای ما آنقدر کم اهمیت است که ایندکس شدن آن سودی برایمان ندارد و حتی ممکن است برای سئو سایتمان هم مضر باشد! گاهی هم بعضی از محتواهایمان با هدف کسب رتبه و دریافت ورودی از گوگل آماده نمیشوند.

مثلاً محتوای صفحه پروفایل کاربری را در نظر بگیرید؛ این صفحه صرفاً بـه درد صاحب آن حساب میخورد و کارایی دیگری ندارد. چگونه گوگل را مجبور کنیم کـه دست از سـر این صفحات یا کلا دست از سر بررسی سایت ما بردارد؟

برای این کار ما میتوانیم <mark>از یک تابلو قرمز رنگ ورود ممنوع به نام «تابلو ایندکسش نکن» (No index TAG) یا تگ «سوپر ایندکسش نکن» (x-robots)، برای ایندکس نشدن صفحات سایت به ره ببریم. با استفاده از ایـن تـگ رباتهای گوگـل متوجـه خواهند شد **این صفحه از خط قرمزهای سایت ماست و نباید وارد آن شود!** امیدوارم تا اینجا خسته نشده باشید! تا اینجا شما با ۱۰ مفهوم اصلی، مهم و تاثیرگذار</mark>

در بودجه خزش آشنا شدهاید. فقط یک مفهوم مهـم دیگر را بایـد بشناسـید و بعـد از آن وارد فصل جذابی از این ایبوک خواهیم شد...

پس بیاید تا سراغ مفهومی برویم که بسیاری از متخصصین سئو را گیج کرده است!



بخش ال کنیبالیزیشن: لطفا در نهایت من را انتخاب کنید (:

اگر من و مهندس امین اسماعیلی در یک جلسه پرسش و پاسخ شرکت کنیم و شما درباره بودجه خزش، سؤالی کنید و ما بهصورت همزمان شروع به جواب دادن کنیم چه میشود؟ قطعاً **شما گیج میشوید و نمیدانید کدام یک پاسخ بهتری برای سؤال شماست**؛ ممکن است گاهی توجهتان را به صحبتهای من دهید و گاهی به صحبتهای مهندس اسماعیلی گوش کنید. در اینجا کنیبالیزیشن رخ داده است و **من با مهندس اسماعیلی کنیبالیزیشن شدهام!!!**

اجازه دهید به کمی به عقبتر برگردیم...

اگر من و مهندس اسماعیلی در یک جلسه پرسش و پاسخ شرکت کنیم و شما درباره بودجه خزش، سؤالی کنید و من در پاسخ به شما «کتاب زمستان سرد ایندکس» (همین ایبوکی که در حال مطالعه آن هستید) و آقای اسماعیلی «دوره سئو تکنیکال» را پیشنهاد دهیم، چه اتفاقی خواهد افتاد؟ آیا باز هم ما باهم کنیبالیزیشن شدهایم؟ خیر! زیـرا ما دو پاسخ متفاوت را بـرای درک مفهوم بودجه خزش بـه شما پیشنهاد دادهایم و دیگر شما در انتخاب گیج نخواهید شد؛ پاسخ به شما کاملاً شفاف است: اگر به دنبال کتاب باشید، من را انتخاب میکنید و اگر محتوای ویدئویی کاملی را با مثالهای عملی نیاز دارید قطعاً مهندس اسماعیلی و دوره سئو تکنیکال را انتخاب

کنیبالیزیشن زمانی اتفاق میافتد که هـدف و مفهوم دو یا چنـد صفحه، بـاهم یکسـان باشند و معمولا در نتایج جستجو باهم رقیب میشوند.

کنیبالیزیشن چیست؟ آموزش شناسایی و رفع مشکل



60

✻

تا اینجا ما سعی کردیم تمام مفاهیم مربوط به بودجه خزش را با زبانی ساده اما حرفهای بررسی کنیم؛ مفاهیمی که تا اینجای کار با آنها آشنا شدید مفاهیم بسیار مهمی هستند و در فصلهای آینده خیلی با آنها سروکار داریم. در این فصل از کتاب ما بهصورت خلاصه با هر مفهوم آشنا شدیم؛ بنابراین پیشنهاد میکنم اگر مفهومی را متوجه نشدهاید یا نیاز به توضیحات بیشتری دارید، حتماً مقالههای کامل آن را در سایت آکادمی وبسیما مطالعه کنید. در سایت آکادمی وبسیما ما به صورت عمیق مفاهیم را بررسی میکنیم و مثالهای بسیار زیادی را برای فهم بهتر مسائل نوشتهایم.

الان وقت آن است که به سراغ آشنایی با ابزارها و نحوه کار با آنها برویم.









ما نمیتوانم به گذشته برگردیم و شروع را تغییر دهیم، بلکه میتوانیم از همینجایی که هستیم، آغاز کنیم و پایان را جور دیگری رقم بزنیم؛ نمیتوانیم در روز اول، سایت بی عیب و نقصی را آماده کنیم؛ اما میتوانیم در طول مسیر، ایرادها و مشکلات آن را رفع کنیم و به نتایج موفقی برسیم. برای این کار از گوگل سرچ کنسول کمک میگیریم. اگر حافظه خوبی داشته باشید، قبلاً به خانم دکتر اشارههایی کرده بودم و شما از دور با آن آشنا شدید؛ حالا زمان آن رسیده که به صورت حرفهای با این ابزار آشنا شوید و قابلیتهای آن را بشناسید. گوگل سرچ کنسول اصلیترین، مهمترین و کاملترین جعبه ابزاری است که ما به صورت رایگان به آن دسترسی داریم و هر چیزی که نیاز

ابزار گوگل سرچ کنسول بخشهای کاربردی متنوعی دارد و اگر بخواهیم به صورت کامل و جامع هر یک از بخشهای آن را بررسی کنیم، نیاز به تدوین یک دوره چند ساعته جامع، همراه با مثالهای عملی داریم. ما در این فصل قصد داریم فقط بخش **گزارشدهی آمار خزش (Crawl Stats)** را بررسی کنیم و همه جوانب را آموزش دهیم. نکته: **البته این دوره، همین الان هم در سایت آکادمی وبسیما در دسترس است.**



خزیدن رباتهای گوگل اصلاً بیحساب و کتاب هم نیست، هر رباتی باید در پایان هر روز، گـزارش کـاملی از فعـالیتش را بـه گوگـل بدهـد. بـه چـه سـایتهای سـر زده، چـه صفحاتی را دیده، با چه پیامها و خطاهـایی مواجـه شـده، چقـدر از منـابع استفاده کرده است و تمام ریز و درشت کارش را گزارش میدهد. (گوگل هم بر اساس همین گزارش میزان حقوق هر ربات را تعیین میکند!)

خبر خوب برای ما چیست؟ <mark>گوگل تمام این گزارشها را به صورت کاملاً شفاف و رایگان</mark> در اختیار ما قرار میدهد</mark> تا با استفاده از آن، بخشی از تکنیکال سایتمان را اصلاح کنیم، با خطاها و مشکلات سایتمان آشنا شویم، بودجه خزش سایتمان را مدیریت کنیم و در نتیجه این فعالیتها، سرعت ایندکس سایتمان را افزایش دهیم.

برای دسترسی به این گزارش ارزشمند **باید از قبل سایت خود را در گوگل سرچ کنسول ثبت کرده باشید** (راهنمای ثبت سایت در گوگل سرچ کنسول) و سپس مراحل زیر را دنبال کنید:

- ۱. وارد ابزار گوگل سرچ کنسول شوید (لینک مستقیم ابزار)
- ۲. از نوار کناری صفحه، وارد گزینه تنظیمات (Settings) شوید
- ۳. در بخش خزیدن (Crawling) گزینه آمار خزش (Crawl Stats) را پیدا کنید
 - ۴. روی باز کردن گزارش (OPEN REPORT) کلیک کنید

· · ·			
Core Web Vitals	Ownership verification	You are not a verified owner	
Enhancements	Associations	Associate a service with this property	
Breadcrumbs	Change of address	Inform Google when your site moves to a new address	
Comparisable structured d	Crawling		
Legacy tools and reports	Crawl stats	4.58K crawl requests (last 90 days)	OPEN REPORT
, Links	About		
\$ Settings	Indexing crawler	Googlebot Desktop	Learn more
Submit feedback About Search Console	Property added to account	July 3, 2022	REMOVE PROPERTY
vacy Terms			





وقتی که این گزارش را باز میکنیم با چنین صفحهای روبرو خواهیم شد:



در ایـن ادامـه قصـد داریـم تـک تـک عناصـر ایـن گـزارش را کالبـد شـکافی و هـر یـک از بخشهای آن را به صورت کامل بررسی کنیم:



requests (بخش)

از نظر مـن نمودار آبـی رنگ، مهمترین بخـش داسـتان آمار خـزش اسـت، ایـن قسـمت تعداد درخواستهایی را نشان میدهدکه خزندههای گوگل در یک بازه زمانی مشخص به منابع سایت ما دادهاند

در تصویر قبل تعداد کل درخواستها برابر با ۴۴۱۰ عدد در بازه زمانی ۳ماهه است. ما میتوانیم این آمار را به صورت روزانه نیز در نمودار مشاهده و بررسی کنیم. چندین نکته مهم برای تحلیل این آمار خزش وجود دارد که اگر این نکات حساس را ندانیم نمیتوانیم گزارش را به صورت درست و دقیق بررسی کنیم:

فقط آمار صفحات اصلي محاسبه ميشوند

تعداد آماری که در این گزارش ارائه میشود **فقط گزارش مربوط به صفحات اصلی سایت ما هستند**؛ اگر یکی از صفحات ما دارای ویژگی تگ کنونیکال به یک نسخه بهتر از خودش باشد یا صفحهای قابل ایندکس نباشد (non index abale) آمار آن در این گزارش محاسبه نخواهد شد؛ بنابراین تعداد مراجعات گوگل خیلی بیشتر از این عدد است.

نکته: گزارش آمار خزش صفحات دارای ویژگی تگ کنونیکال در قسمتهای دیگر گوگل سرچ کنسول قابل بررسی و تجزیه و تحلیل است.

ACADEMY

هر حرکت، یک درخواست

همین امروز صبح یک اسنپ گرفتم تا به یک جلسه کاری مهم بروم، اما تقریباً در نیمههای راه، جلسه لغو شد و مجبور به انتخاب مقصد دوم شدم و به محل کارم برگشتم؛ به محل کارم که رسیدم، راننده اسنپ هر دو کرایه مقصد اول و دوم را از من گرفت و وقتی علتش را پرسیدم، به من گفت: «من ربات گوگلم؛ حقوق گوگل کفاف زندگی رو نمیده مجبورم کنار کار، اسنپ هم برم! توی گوگل وقتی شما یه آدرس رو ریدایرکت کرده باشی **ما ۲ تا درخواست میدیم**، یکی برای صفحه ریدایرکت شده (۳۰۱) و یکی هم برای صفحه اصلی (۲۰۰)، الانم شما آدرست رو ریدایرکت کردی، منم ۲ تا پول میگیرم ازت...»

خوب میدانم که شما این داستان کاملاً واقعی را باور نکردهاید، شاید باورش برایتان سخت باشد؛ **اما این داستان کاملاً واقعی بود!** (البته قسمت دومش) <mark>گوگل به میزان</mark> تعداد ریدایرکتها، درخواست ارسال میکند و آمار آن، در این قسمت هم خواهد آمد. مثلاً اگر در سایتمان یک زنجیره ریـدایرکت داشـته باشـیم، مثلاً چنـد صفحه را روی یکدیگر ریدایرکت کرده باشیم، بینهایت هم درخواست برای ما ثبت میشود. **البتـه هرگـز ایـن اتفـاق نخواهـد افتـاد**؛ زیـرا رباتهـای گوگـل بـا چنـد بـار دور خودشـان چرخیدن متوجه خواهند شـد که شـما آنهـا را سر کار گذاشـتهاید و میرونـد تا گـزارش

شما را بدهند ... فكر نمىكنم اين اتفاق خوبى باشد.



اگر پولهایتان را بشمارید، کم میشود

ما بعضی از صفحات سایت را با استفاده از فایل دستورالعمل رباتها (robot.txt)، محدود کردهایم و به رباتهای گوگل فهماندهایم این صفحات را بررسی نکنید و اگر هم فضولیتان گل کرد و به آن سری زدید لطفاً به مسئول ایندکس نگویید تا ایندکس نشوند. نکته باحال ماجرا دقیقاً اینجاست که **آمار فضولیهای رباتهای گوگل هم در این آمار گزارش میشود**.

میزان خزیدن صفحات موجود در فایل دستورالعمل رباتها، در آمار کلی بالای صفحه محاسبه میشود؛ اما به جدولهای تفکیکی پایین صفحه اضافه نمیشود؛ اگر دوست دارید آمار میزان فضولیت رباتهای گوگل را محاسبه کنید، باید آمار یکی از جدولهای تفکیکی پایین را با هم جمع و از آمار کلی بالا کم کنید؛ بنابراین آمارهای تفکیکی عدد کمتری نسبت به آمار کلی دارند و ممکن است باهم مساوی نباشند.

گوگل آدم فروش نیست

باید در سایت یکی از دوستانم تصاویر بسیار باکیفیت و با حجم بالایی قرار میدادیم و به همین دلیل هم به سرور آن فشار وارد میشد و هزینه یک سرور عالی برای کل سایت خیلی گران میشد، تصمیم گرفتیم یک زیر دامنه برایش بسازیم و تصاویر را آنجا آپلود کنیم، سپس یک سرور قوی کوچک را برای آن فعال کنیم و سرور سایت اصلی را تغییر ندهیم.

در نتیجه در هزینهها صرفه جویی کردیم و مشکلی هم در سایت ایجاد نشد. پس از چند وقت دوستم از من بسیار شاکی شد و گفت که پولم هدر رفت و کاری که کردیم به سایتم آسیبی جدی زده است. وقتی پیگیر داستان شدم گفت که **گوگل به هیچ عنوان تصاویر سایت را نخزیده است** و این یعنی فاجعه!

اما اتفاق هولناک این بود که نمیدانستیم **باید در سرچ کنسول زیر دامنهای وارد شویم که تصاویر را در آن آپلود کردهایم**؛ مـن میدیـدم گوگل، تصاویر را اینـدکس کـرده و در نتایج هست و همـین موضوع برایمـان عجیب بود تـا اینکـه بعـد از بررسـیهای عمیق، متوجه این نکته ریز شدیم و مشکل حل شد.

گوگل فقط و فقط آمارهای مربوط به یک سایت و یک دامنه را به ما ارائه میدهد و اگر منابعی (مانند فونتها، جاوا اسکریپت، تصاویر و ...) از سایتها یا حتی زیر دامنههای دیگری در صفحات سایت ما فراخوان شوند، **به هیچ عنوان در گزارش خزیدن سایت و دامنه ما جایی نخواهند داشت**؛ بنابراین اگر نرخ خزش بعضی از منابع مانند تصاویر (بهشرطی که از سایت دیگری فراخوان شوند) پایین بود تعجب نکنید.

نکته: برای بررسی آمار خزش منابع فراخوان شده، باید وارد سرچ کنسول همان سایت شویم و یا اگر از زیر دامنههای خودمان فراخوان میشوند، باید با تعریف کردن سطوح دامنه (domain level property)، به آمارهای آن دسترسی داشته باشیم.

گوگل دوبارہ کاری نمیکند

گوگل بعضی از منابع سایتها مثل جاوا اسکریپتها، تصاویر، سی اس اسها را برای خـود ذخیـره (کـش) کـرده تـا در منـابع مـا و خـودش صـرفهجویی کنـد؛ یـک فایـل استایل (css) ممکن است در تمامی صفحات ما استفاده شده باشد و این به هیچ عنوان منطقی نیست که رباتها به هر صفحهای که سر زدند، آن فایل را بررسی کنند. اگر شما از یک تصویر در چندین صفحه سایتتان استفاده کردید، فقط یکبار خزش و محاسبه میشود.



نکته: به مرور زمان درصد این گونه فایلها کاهش مییابد و به همین دلیل هم در گزارش تفکیک شده در زیر گزارش، درصد این گونه فایلها بسیار پایین است؛ زیرا قبلاً خزیده و کش شدهاند.

درک بعضیها سخت است

بعضی از زن و شوهرها را دیدهایم که پس از یک عمر زندگی از یکدیگر جدا میشوند، اگر دلیل این اتفاق را جویا شویم ممکن است بعضی از آنان بگویند که «**من تازه او را شناختهام!**» اما بعضی از آدمها را هم میشود از همان برخورد اول شناخت. این قانون در گزارش آمار خزش هم صدق میکند.

تعداد این درخواستها با تعداد صفحات سایت ما برابر نیست؛ هر فایل جدیدی (مثل یک عکس) یک درخواست مجزا محاسبه میشود و ممکن است یکی از صفحات نیاز به ۱۰۰ بار درخواست خزش داشته باشد.

مثلاً صفحه لندینگی را در نظر بگیرید که کاملاً اختصاصی طراحی شده است و ۵۰ تا عکس و آیکون جدید در آن وجود داشته باشد. گوگل که قبلاً این تصاویر و آیکونها و طراحی اختصاصی ما را ندیده است؛ بنابراین باید **بیش از ۵۰ درخواست خزش** برای تصاویر و برای سایر منابع لندینگ بدهد. در نتیجه تعداد این درخواستها برابر با تعداد صفحات ایندکس شده نیست و **منظور از تعداد کل خزش، فقط مربوط به تعداد** د**رخواستهاست** و ربطی به تعداد صفحات ندارد.

نکته: در این گزارش لیستی از صفحات را نیز مشاهده میکنیم که این لیست تمام صفحات خزیـده شـده نیسـت و **فقـط نماینـدهای از ویژگیهـای آنهـا هسـتند**؛ سـخنی از نویسنده: «اگه صفحهای در گزارش نیست، هستش و به معنی کلاً نیستش نیست!»



هر چیزی کم یا زیادش لزوماً خوب نیست

اگر به نمودار روزانه نگاهی بیندازید، متوجه میشوید در بعضی از روزها به شدت رشد کرده و کمی بعد افت شدیدی داشته است. این **تغییرات ناگهانی لزوماً از وجود یک مشکل خبر نمیدهد** و ممکن است دلیل آن ایجاد تغییراتی باشد که روی سایت اعمال کردهایم.

اگـر ۱۰ هـزار مخاطـب داشـته باشـیم و بـا اسـتفاده از فایـل دسـتورالعمل رباتهـا (robot.txt) صفحه پروفایـل کـاربری مخاطبینمـان را مسـدود کنیم، شـاهد یـک افـت ناگهانی در این نمودار خواهیم بود. یا برعکس اگـر نقشـه سـایت بهروزرسـانی شـده را بـه گوگـل معرفـی کنیم، رباتهـا آدرسهـای جدیـد نقشـه سـایت را بررسـی میکننـد و ایـن نمودار، رشد قابل توجهی خواهد کرد.

نکته: رشد و افت ناگهانی این نمودار فقط در صورتی خطرناک است که تغییری را اعمال نکرده باشیم. مثلاً علت افزایش و کاهش آن به ترتیب میتواند این ۲ مورد باشد:

- سایتمان هک شده و ۱۰ هزار صفحه جدید روی آن ساخته شده است.
 - سرورمان به مشکل خورده و از کار افتاده است

امنیت مهم نیست

این گزارش تفاوتی بین صفحات HTTP و HTTPs قائل نمیشود و گزارش هر دو را یکجا میتوانیم مشاهده کنیم؛ بنابراین نیازی به بررسی جداگانه هر کدام نداریم. نکات مهم این قسمت از گزارش تمام شد و الان باید سراغ بخش دوم برویم.



بخش Total download size (Bytes) بخش ۲

این قسمت حجم کل فایلها و منابع را در یک بازه زمانی مشخص و بر اساس واحد بایت نشان میدهد که در طول فرایند خزش دانلود شدهاند؛ در تصویر بالا حجم کل فایلهای دانلودی برابر با ۱۲۳ میلیون بایت (۱۲۳ مگابایت) بوده است. نکته: همانطور که قبلاً هم بررسی کردیم، اگر فایلی کش شده باشد دیگر برای آن

درخواستی داده نمیشود و بنابراین دانلودی هم اتفاق نخواهد افتاد.

Average response time (ms)

میانگین زمان پاسخگویی، بـه هـیچ عنـوان مـدت زمـان بارگـذاری صـفحه نیسـت! ایـن قسـمت فقـط و فقـط مـدت زمـانی را محاسـبه میکنـد کـه طـول میکشـد تـا سـرور بـه درخواست خزنده پاسخ دهد.

توجه داشته باشید در این زمان هیچ منبعی (css، تصاویر، جاوا اسکریپت و...) و **حتی** HTML **صفحه نیز بارگذاری نشده است**. (دقیقاً همان زمانی که دوستانتان را صدا میزنید و مثلاً ۲ ثانیه بعد او به شما نگاه میکند، هنوز پاسخی به شما نداده و حرفی نزده است.)

نکتـه: <mark>ایـن سـرعت بـه مـواردی مثـل موقعیـت (Location) دی ان اس، سـی دی ان،</mark> کشـینگ، اس اس ال، رم و سـی پـی یـو سـرور وابسـته اسـت پاسخگویی بالاتری داشته باشیم، یک نکته مثبت برای ماست و روی فاکتورهـای هسته حیاتی سایت (core web vitals) تاثیرگذار است.

تا اینجا شما کاملاً با قسمت بالای نمودار و کمی هم با قسمت پایین نمودار آشنا شدید و قصد داریم در ادامه به صورت کامل، بخشهای تفکیکی پایین نمودار را نیز بررسی کنیم. این قسمت بر اساس چند ویژگی، گزارش خزش را دستهبندی میکند:

- .۱ احوالات صفحه (By response).
 - ۲. فرمت منابع (By file type)
 - ۳. هدف بررسی (By purpose)
- ۴. ربات مسئول (By Googlebot type)

By response			By file type		
OK (200)	89%	house	HTML	57%	und
Not modified (304)	5%		Image	16%	self-han
Moved permanently (301)	3%		JavaScript	7%	<u> </u>
Not found (404)	2%		Syndication	6%	
Page could not be reached	< 1%		CSS	3%	
Rows per page: 5 👻	1-5 of 7	< >	Rows per page: 5 👻	1-5 of 10	< >
By purpose			By Googlebot type		
By purpose Refresh	84%	himmedia	By Googlebot type Smartphone	49%	-
By purpose Refresh Discovery	84%	harmale.	By Googlebot type Smartphone Image	49% 21%	مهرسان بطوراری
By purpose Refresh Discovery Rows per page: 5 💌 1	84% 16%	himmelie. Literaturan K	By Googlebot type Smartphone Image Desktop	49% 21% 18%	No
By purpose Refresh Discovery Rows per page: 5 👻 1	84% 16% 1-2 of 2	himmelie. Literaturan K	By Googlebot type Smartphone Image Desktop Page resource load	49% 21% 18% 11%	مهرددار بطوراری استور
By purpose Refresh Discovery Rows per page: 5 🛩 1	84% 16% 1-2 of 2	himmelie. Libertum < >	By Googlebot type Smartphone Image Desktop Page resource load AdsBot	49% 21% 18% 11% 1%	Na



بخش ۳ احوالات صفحه (By response)

در این قسمت خانم دکتر گوگل سرچ کنسول، احوال صفحاتی را به صورت تفکیک شده گزارش میدهد که آنها را معاینه کرده است. مثلاً وضعیت «اوکی» (۲۰۰) یعنی صفحه سالم بود و برای من کاملاً باز شد یا وضعیت «پیدا نشد» (Not Found) به ما میگوید که این صفحه علائم حیاتی ندارد و دیگر امیدی به آن نیست.

آیا همه وضعیتها هستند؟

در این قسمت ممکن است وضعیتهای دیگری را نیز ببینید؛ یادم میآید آقای پشمک از مهندس اسماعیلی سؤال میکرد که چرا این قسمت از گزارش سایت من، وضعیت ارور سرور (Server error) را ندارد! خب مثل این است که آقای پشمک سرما خورده باشد و به دکترش بگوید: «چرا فقط قرص سرماخوردگی تجویز کردی و در نسخه من آمپول هاری نیست؟» **البته شاید هم سرما خورده و هم سگ گازش گرفته است!** اگر سایر وضعیتها مانند عدم وجود فایل دستورالعمل (robots.txt not available) را ندارید، یعنی سایت شما این خطا را ندارد و نیازی نیست نگران آن باشید.

بررسی سلامت سایت یا یک نگاہ

این نکته را یادتان نرود: «تنها با یک نگاه به آمار این قسمت، میشود وضعیت سلامت سایت را تشخیص داد. این کار از پایه و اساس اشتباه است و **هرگز** نمیتوان با یک نگاه سرسری متوجه وضعیت سلامت سایت شویم.» (باز هم غافلگیرتان کردم!) تک تک این وضعیتها و عناوین را باید با دقت بررسی کنیم؛ به آمار بالا دقت کنید، میزان درصد صفحات ۴۰۴، ۲ درصد است، **اگر همین ۲ درصد تعدادی از مهمترین صفحات سایت ما باشند چه؟** یا اگر ۲ درصد از ۱ میلیون صفحه (۲۰ هزار تا) باشد چه؟ بنابراین نباید گول این عددها و درصدها را بخوریم.

اگر این درصدها کارایی ندارند پس چرا هستند؟ اینطور هم نیست! با یک نگاه به این نمودارها هم میتوان موارد مهمی را متوجه شد؛ باید به روند صعود یا نزول این نمودار دقت کنیم. مثلاً اگر آمار وضعیت صفحات اوکی (۲۰۰) یک رشد ناگهانی را داشته باشد باید به دنبال دلیل این اتفاق باشیم. **در وضعیت بد شاید سایتمان هک شده یا تیم فنی جایی را سهل انگاری کرده است** و در وضعیت خوب هم شاید هم خودمان تغییراتی را ایجاد کرده باشیم. (مثلاً نقشه سایت را بهبود دادهایم)

بخش ۴ فرمت منابع (By file type)

همان مثال خانم دکتر را در نظر بگیرید، او در این قسمت آمار نوع صفحات را مشخص میکند؛ مثلاً مراجعهکنندگانش، زن بودند یا مرد، جوان بودند یا پیر، کودک بودند یا خردسال؟ بنابراین در این قسمت میتوانم بر اساس فرمت منابعی که بیشترین مصرف را داشتهاند گزارش را تجزیه و تحلیل کنیم. مثلاً متوجه میشویم تصاویر سایتمان چند درصد از میزان خزش را به خود اختصاص دادهاند. نکته: فرمتهای HTML همان آدرسهای اصلی ما هستند.

بخش هدف بررسی (By purpose)

قبلاً در فصل دوم زمانبندی خزیدن را بررسی کردیم: «اگر گوگل صفحهای را بررسی کند و متوجه شود محتوای آن تغییری نمیکند، کمتر آن را میخزد. برای مثال اگر گوگل یک صفحه را بخزد و در روز بعد هیچ تغییری را در آن نبیند، ممکن است ۳ روز بعد دوباره به سراغ آن بیاید و اگر باز هم تغییری نبیند ۱۰ روز دیگر دوباره آن را بررسی کند و همینطور بازههای زمانی بررسی را به ۳۰ – ۵۰ – ۱۰۰ و بیشتر به تأخیر بیندازد.» در این بخش ما دو آمار متفاوت «تازه» و «کشف» را دایم؛ در یک نگاه شاید تفاوتی نداشته باشند؛ اما در عمل از زمین تا آسمان تفاوت دارند.

تازگی (Refresh)

این قسمت «مادر بزرگ، خانمجان» **کلییییید اسرار!!!!** (نکته: اگر سریال کلید اسرار را ندیدهاید پیشنهاد میکنم حتماً یکی از قسمتهای آن را ببینید!!!)

مادر بزرگ «خانمجان» در یکی از بیمارستانهای قلب شیراز بستری است و دکترها به او گفتهاند که تا چند روز دیگر بیشتر زنده نیست. در یکی از این روزها عزرائیل بـه سـراغ او میآید و میگوید: «فردا شب آماده باش که باید باهم برویم...»، برای همین هم خانم جان تصمیم میگیرد برای امور ساخت مدرسه و بیمارستان در شهرهای حاشیهای کشور، اموالش را به خیریه اهدا کند. فردای آن روز دکترها با کمال تعجب میبینند که او کاملاً سر حال و سلامت است و شب هنگام، عزرائیل میآید و به او میگوید: «تو به دلیل کار خوبی که کردی ۵۰ سال به عمرت اضافه شد و تا ۱۰۸ سالگی میتوانی زنده باشی» خانم جان با شنیدن این خبر خوشحال میشود و فردای آن روز، تمام عملهای زیبایی را کـه میتوانسـت و نمیتوانسـت را انجـام داد و اسـمش را از «خانمجـان» بـه «پارمیدا جون» عوض کرد. وقتی از بیمارستان مرخص شد و هوای تازه را نفس کشید با صدای بلند گفت: «زندگی جدید **سلااااام!**» دقیقا در همین لحظه یک پراید دور رنگ مدل ۸۴ او را زیر گرفت و جان به جان آفرین تسلیم کرد. وقتی دلیل این موضوع را از عزرائیل پیگیر شد، عزرائیل به او گفت: «**عه عه عه!** خانم جان توپیی؟؟؟ تغییراتت اصلاً ایندکس نشده بود!!! نشناختمت...»

گوگل با درس گرفتن از این داستان متوجه شد باید مدام به صفحات سایتها سر بزند تا اگر تغییری داشتهاند آنها را در فهرست ایندکس خود به روز کند و خدایی ناکرده باعث و بانی اتفاق بدی نباشد!!! مثلاً با اتمام مدت زمان تخفیف شب یلدا، تخفیف را از عنوان حذف میکنیم و اگر گوگل آن را دوباره ایندکس نکند، در نتایج همراه با عبارت

«تخفیف ویژه ۵۰ درصدی» دیده میشویم و باید تا شب عید نوروز تخفیف بدهیم. ۲۲ ۸۲ ACADEMY اگر گوگل صفحهای را بررسی کند و متوجه تغییراتی در آن شود، باید آن را دانلود و جـایگزین نسـخه قبلـی کنـد</mark>. آمـار چنـین صـفحاتی را در قسـمت تـازگی (Refresh) میتوانیم بررسی کنیم. مثلاً در نمودار ۸۴ درصد از صفحات ما از این جـنس هستند و مابقی «کشف» (Discovery) هستند.

كشف (Discovery)

اگر رباتهای گوگل به صفحهای برسند که قبلاً آن را بررسی نکردهاند (مثلاً مقالـهای کـه امروز آن را منتشـر کـردهایم) ایـن یـک صفحه جدیـد است و آمـار آن را در بخـش جـدایی ثبت میکنند.

چند نکته مهم:

۱. میزان درصد این دو فاکتور به میزان نرخ تولید محتوای سایت بستگی دارد

۲. اگر سایتی تازه تأسیسی داریم قطعاً قسمت دیسکاوری آن بیشتر از رفرش است

۳. برای بررسی این قسمت باید به ماهیت سایت توجه کنید (مثلاً سایتهای خبری باید درصد دیسکاوری بیشتری داشته باشند)

۴. همچنان گول درصدها را نخورید و هر دو را کاملاً بررسی کنید

۵. تغییرات کوچک و بیارزش در سایت مثلاً یک تغییر کوچک در (CSS) نرخ رفرش را

بالا میبرد و این باعث نابودی بودجه خزش میشود؛ یعنی به جای اینکه بودجه خزش

صرف محتواهای جدید و ارزشمند ما را شوند، صرف بازبینی صفحات قبلی میشوند



بخش ۶ ر**بات مسئول (By Googlebot type**

هر یک از رباتها مسئول بررسی قسمت خاصی از سایت هستند، مـثلاً ربـات Smatphone فقط نسخه موبایل سایت را بررسی میخزد و ربات image، فقط تصاویر را بررسی میکند. در این قسمت میتوانیم میزان خزش را بر اساس نوع رباتها تجزیه و تحلیل کنیم. اگر به آمار نگاه کنید، درصد ربات اسمارت فون، ۴۹ درصد و درصد ربات دسکتاپ ۱۸ درصد است؛ و ایـن دقیقاً الگوریتم «Mobile First Index» را نشـان میدهد.

الگوریتم Mobile First Index و سفر به اعماق زمین

در این فصل به صورت کامل و جامع، گزارش آمار خزش (Crawl Stats) را در ابزار گوگل سرچ کنسول بررسی کردیم و در حال حاضر زمان آن رسیده تا آستینهایمان را بالا بزنیم و شروع به بهینهسازی بودجه خزش کنیم.



9

✻





موارد متعددی جوود دارد که ممکن است گوگل به دلیل وجود آنها وبسایت ما را به انـدازه کـافی نمیخـزد. در ایـن فصـل قصـد داریـم بهصـورت کامـل و مرحلهبهمرحلـه، اصلیترین و مهمتـرین راهکارهـای عملی را بـرای افـزایش بودجـه خـزش بررسی و اجـرا کنیمو به صورت مرحله به مرحله اقدامات لازم برای اجرای آنها را طی کنیم. اگر آماده هستید تا با ۱۴ ترفند و تکنیک فوق العـاده موثر و جـذاب آشنا شوید و برای خود از آن استفاده کنید، سریع به سراغ مطالب این فصل میرویم.





^{بخش ا} خزیدن تو، تکراری است که تکراری نمیشود

صفحاتی که از لحاظ محتوا و ساختار دقیقاً کپی یکدیگر یا بسیار شبیه به هم باشند باعث ایجاد محتواهای تکراری در سایت میشوند. وجود این محتواها باعث میشود که **گوگل، صفحات کپی را از صفحه اصلی مدنظرمان تشخیص ندهد** و در نتیجه این اتفاق، سرعت ایندکس ما با کاهش چشمگیری مواجه شود.

استفاده درست و اصولی از تگها

بسیاری از افراد درک درستی از کاربرد تگ (هشتگ) در سایت ندارند و مانند شبکههای اجتماعی از آن استفاده میکنند؛ برای مثال من خودم صفحههای زیادی را دیدهام که از هشتگهایی بیهدف و به صورت غیر اصولی استفاده کردهاند و تمام عبارات کلیدی خود را در زیر صفحه هشتگ زدهاند.

اگر کاربرد درست هشتگها را ندانیم، این تگها باعث ایجاد محتواهای بیارزش و تکراری زیادی در سایت ما میشوند؛ زیرا هر هشتگ یک صفحه در سایت ما میسازد! در دوره تدوین استراتژی سئو ما قبل از شروع فعالیت سایت، هشتگهای خود را به درستی انتخاب و از آنها استفاده میکنیم. اگر در سایت خود به صورت غیر اصولی از هشتگها استفاده کردهاید، باید صفحات مشابه را پیدا و آنها را روی موردنظر ریدایرکت کنید.

تدوین استراتژی سئو؛ مراحل ساخت یک امپراطوری پایدار



☆

کمکگیری از لینکسازی داخلی

لینکسازی داخلی در نشاندادن صفحه درست به موتورهای جستجو کمک فراوانی میکند. اگر میخواهیم گوگل در یک عبارت کلیدی (مثلاً خرید عطر بیک)، صفحه مشخصی را بهعنوان صفحه هدف ما بشناسد، باید برای آن لینکسازیهای داخلی حرفهای انجام دهیم. یعنی از صفحات دیگر سایت خود (بهویژه صفحات مهم سایتمان)، به صفحه هدف لینک دهیم. **انجام این کار باعث میشود که گوگل اهمیت** صفحه هدف را متوجه شود.

در دوره لینکسازی درباره تمام اصول لینکسازی صحبت میکنیم و از همه جوانب، اصول لینکسازی داخلی و خارجی را بررسی کردهایم.

جلوگیری از ایندکس محتواها بر اساس فرمت

وردپرس با تمام خوبیهایی که دارد، اگر حواسمان به کارهایش نباشد و کنترلش نکنیم، ممکن است شیطنتهایی جبرانناپذیر کند. <mark>یکی از ایرادات وردپرس این است</mark> که محتواهای سایت ما را بر اساس فرمت، آرشیو میکند و همین باعث هدر رفتن بودجه خزش و ایجاد محتواهای تکراری میشود.

برای مثال وقتی ما یک عکس یا ویدئو را در مقاله خود قرار میدهیم، آن عکس یا ویدئو علاوه بر خود مقاله، در فضای آرشیو هم ذخیره میشود و یک لینک جدید را به خود اختصاص میدهد. این مسئله باعث ایجاد محتوای تکراری شده و در نتیجه آن، بودجه خزش و سرعت ایندکس شما کاهش پیدا میکند.

افزونه Yoast SEO

برای حـل ایـن مشـکل کـافی اسـت پـس از نصـب ایـن افزونـه وارد بخـش تنظیمـات آن شویم و گزینه ایندکسشدن رسانهها را غیرفعال کنید.

این گزینه، **فقط روی بخش مربوط به آرشیو آن تأثیرگذار است** و به معنی عدم ایندکس تصاویر شما نیست.

نکته: در اکثر موارد، زمانی که شما Yoast SEO را نصب میکنید، خود افزونه بهصورت اتوماتیک این گزینه را غیرفعال میکند و لازم نیست شما این کار را انجام دهید.

بخش ۲ ساخت صفحات موبایل-فرندلی (Mobile-Friendly)

از سال ۲۰۱۹ به بعد، گوگل تغییرات بسیار بزرگی در عملکرد رباتها و نحوه خزش سایتها به وجود آورد. با آپدیتی که در سال ۲۰۱۹ انجام شد، **گوگل بخش موبایل وبسایت را بهطورکلی از بخش دسکتاپ جدا کرد** و اهمیت وجود نسخه موبایل بسیار بالا رفت؛ بهطوری که امروزه رباتهای موبایل گوگل، در ابتدا نسخه موبایل صفحات سایت ما را بررسی میکنند.

در نتیجه، یکی از فعالیتهایی که نرخ خزش ما را بهشدت بالا میبرد، ساخت صفحات موبایل فرندلی است. اگر صفحات موبایل سایت ما باکیفیت و سرعت بارگذاری بالایی داشته باشند و نیازهای کاربران را برطرف کنند، نرخ خزش سایت ما افزایش خواهد یافت؛ بنابراین ما باید تمام تلاش خود را بکنیم تا بهترین نسخه موبایل را ارائه دهیم.

استفاده از google page insights

این ابزار محبوب گوگل به ما کمک میکند تا نسخه موبایل صفحات خود را بهینهسازی کنیم و کیفیت آنها را بهبود دهیم. برای استفاده از این ابزار، مراحل زیر را دنبال کنید: ۱. وارد وبسـایت <u>pagespeed.web.dev</u> شـوید. (ممکـن اسـت بـا آیپـی ایـران در دسترس نباشد)

۲. آدرس صفحه موردنظر خود را وارد کنید. ۳. روی گزینه Analyze کلیک کنید و منتظر نتایج بمانید.

۴. بعد از نشان داده شدن نتایج، دو گزینه موبایل و دسکتاپ در دسترس است. گزینـه موبایل را انتخاب کنید. (در بیشتر مواقع نسخه موبایل مهمتر است)

۵. در این قسمت، این ابزار به کیفیت صفحه سایت ما در نسخه موبایل، رتبهای بین ۰ تا ۱۰۰ میدهد. در پایین صفحه نیز تمامی اطلاعات، ایرادها و توضیحات لازم نیز برای ارتقا و بهبود صفحه موبایل قابل مشاهده است. اگر از لحاظ فنی توانایی انجام اصلاحات و اقدامات لازم را دارید، طبق اطلاعات و توضیحات داده شده در این ابزار اقدام کنید؛ در غیر این صورت میتوانید از یک متخصص بخواهید تا اصلاحات را انجام دهد. (در دوره سئو تکنیکال، انجام اصلاحات این قسمت را در دو بخش تئوری و عملی آموزش دادهایم)

۶. شما با انجام اصلاحات لازم و بهبود کیفیت صفحات وبسایت در نسخه موبایل، نرخ خزش و سرعت ایندکس سایت خود را افزایش دادهاید.

استفاده از بخش inspect مرورگر

یکی از راههای دیگر بـرای بررسـی کیفیـت نسـخه موبایـل صـفحات، اسـتفاده از ابـزار inspect مرورگر است. اقدامات زیر را انجـام دهیـد تا بتوانیـد کیفیت و سـرعت صـفحه موردنظر خود را روی موبایل تست کنید.

۱. ابتدا با دسکتاپ، وارد مرورگر شوید و صفحه مورد نظر را باز کنید. ۲. روی یک قسمت از صفحه کلیک راست کنید و گزینه inspect را بزنید. ۳. در تب باز شده روی گزینه Lighthouse کلیک کنید. ۴. شما با یک صفحه مواجـه میشوید کـه دارای آپشـنهایی مختلف اسـت. از قسمت

device، گزینـه موبایـل را انتخـاب و از قسـمت کتگـوری تیـک گزینـه performance را فعال کنید. (برای دسترسی به نتیجهای شفافتر، تیک سایر گزینهها را بردارید.)

۵. روی گزینه Analyze Page Loadکلیک کنید و برای مدتی صبر کنید تا تستهای لازم انجام شود.

۶. از بین • تا ۱۰۰ رتبهای به کیفیت صفحه موبایل شما داده میشود و در ادامه آن، تمام خطاها، ضعفها و اصلاحات قابل انجام برای این صفحه به شما نمایش داده میشود. حال شما میتوانید بهتنهایی یا به کمک یک متخصص برای حل این مسائل و بهبود نسخه موبایل صفحات اقدام کنید.



^{بخش ۳} ارتقای کیفیت و سرعت سرور

همانطور که بارها گفتهایم، **سرعت سرور در میزان بودجه خزش سایت بسیار تأثیرگذار** است. اگر سرور ما کیفیت مطلوب مورد نیاز را نداشته باشد نمیتوانیم به تمامی ورودیهای سایت و درخواستهای رباتها گوگل پاسخ بدهیم. مثلاً با ورود تعداد زیادی کاربر و ربات، وبسایت ما باز نخواهد شد و گوگل نرخ خزش خود را نسبت به سایت شما کاهش میدهند.

استفاده از هاست باکیفیت

ما باید حتماً از یک هاست باکیفیت استفاده کنیم. با وجود هاست باکیفیت، خیالمان از سرعت و کیفیت سرور راحت خواهـد بـود و علاوهبرآن، گوگـل و موتورهـای جسـتجو بـا سرعت بیشتری صفحات مختلف سایت را ایندکس میکنند.

برطرفكردن خطاهاى سرچ كنسول

در نتیجه وجود خطاها، کیفیت سایت و در نهایت بودجه خزش ما کاهش یابد. برای بررسی خطاهای سایت باید وارد ابزار گوگل سـرچ کنسول شویم و ایـن مراحـل را دنبال کنیم: indexing < pages

حال، تمامی صفحات و بخشهای مختلف سایت ما در دسترس هستند و اگر صفحهای دارای خطا باشد در اینجا نمایش داده میشود. در پایین این بخش، دستههای مختلف ارور به شما نشان داده شده است و با کلیک روی هر یک از آنها میتوانید صفحات مربوطه را مشاهده کنید. پس از بررسی این قسمت، حتماً برای رفع ارورها اقدام کنید. انجام این کار به افزایش نرخ خزش سایت ما کمک بسیاری میکند. برای نمونه برخی از این خطاها به این صورت است که در فصل مفاهیم مربوطه درباره آنها مطلبهای



مفیدی نوشتیم:
زمستان سرد ایندکس: فصل اول

- Alternate page with proper canonical tag
- Page with redirect
- Not found (۴∘۴)
- Excluded by 'noindex' tag
- Duplicate without user-selected canonical
- Server error (Δxx)
- Soft ⊬₀۴
- Blocked due to access forbidden (۴۰۳)
- Crawled currently not indexed
- Discovered currently not indexed

^{بخش ۴} افزایش نرخ خزش سایت با استفاده از API گوگل

عبارت نام آشنای «Api» مخفف «application programming interface» است. ابزار indexing api به ما کمک میکند زمانی که یک صفحه یا آدرس جدیدی را به سایت خود اضافه یا حذف کردیم، به موتورهای جستجو درمورد آن اطلاع دهیم. به کمک ایـن کـار مـا میتوانیم سـرعت اینـدکس را بـالا ببـریم و بودجـه خـزش سـایت را افزایش دهیم. indexing api در افزایش بودجه خزش سایت بسیار تأثیرگذار است.

فعالسازی api

برای استفاده از این ابزار باید مراحل زیر را دنبال کنیم. (توجه: ممکن است این ابزار با آیپی ایران در دسترس نباشد)

۱. آدرس <u>https://console.cloud.google.com/apis/library</u> را باز کنید.

۲. وارد اکانـت گوگـل خـود میشـویم. اگـر تابـهحال در پلتفـرم Cloud گوگـل اکانـت نداشتهایم و برای بار اول است که وارد آن شویم، باید کشور خود را انتخـاب کنـیم. برای مثال میتوانیم کشور کانادا یا انگلیس را انتخاب کنیم. (تفاوت زیادی ندارد)

۳. با استفاده از بخش جستجو در بالای صفحه، عبارت Indexing API را جستجو میکنیم.

۴. از بین گزینههای موجود روی Indexing API کلیک میکنیم.

۵. حال کافی است روی گزینه Enable کلیک کنیم. پس از آن، api به راحتی برای ما فعال میشود.

۶. سپس وارد صفحه جدیدی میشویم که تمام دادهها و دریافتیها این api در آن قـرار دارد. اگرچـه چـون بـهتازگی آن را فعـال کـردهایم، هـیچ دادهای بـرای نمـایش بـه مـا وجود ندارد.

۷. روی گزینه Credentials در سمت چپ صفحه (آیکون کلید) کلیک میکنیم.

۸. روی عبـارت Create Credentials در بـالای صـفحه کلیـک مـیکنیم و گزینـه Service Account را انتخاب میکنیم.

۹. در قسمت service account name نام سرویس خود را بهدلخواه وارد میکنیم و در قسمت service account description یک توضیح دلخواه درمورد این سرویس میدهیم.

۱۰. روی گزینه Create & Continue کلیک میکنیم.

۱۱. در باکس role، گزینه owner را انتخاب و روی continue کلیک میکنیم.

۱۲. حال بر روی گزینه Done کلیک میکنیم.

۱۳. تـا ایـن مرحلـه، مـا api خـود را وارد کـردهایم و یـک اکانـت بـرای راهانـدازی آن ساختهایم. حالا میخواهیم این اکانت را به وبسایت خود متصـل کنیم. برای این کار وارد همان قسمت قبلی یعنی Credentials میشویم. میتوانیم در قسمت Service Accounts، اکانتی که بهتازگی ساختهایم را مشاهده کنیم. روی آن کلیک میکنیم.

۱۴.پس از باز شدن اطلاعات اکانت جدید، روی تب keys کلیک میکنیم.

۱۵.گزینه add key را میزنیم و عبارت create new key را انتخاب میکنیم؛ سپس روی گزینه JSON کلیک میکنیم و Create را میزنیم.

۱۶.حال یک فایل json در سیستم کامپیوتر ما دانلود میشود و کار ما در این بخش به اتمام رسیده است. گام بعـدی، متصلکردن api به گوگل سرچ کنسول است که در ادامه آن را انجام میدهیم.

۱۷. نکته: وارد بخش details همان صفحه میشویم. در آن قسمت ایمیل یک آدرس طولانی ایمیل نوشته شده است. آن را کپی میکنیم؛ زیرا برای ثبت در سرچ کنسول به آن نیاز داریم.

اتصال indexing api به سرچ کنسول گوگل

۱. وارد گوگل سرچ کنسول خود میشویم و به بخش setting میرویم. ۲. روی users & permissions کلیک میکنیم و گزینه add user را میزنیم. ۳. ایمیلی قبلاً کپی کردهایم را در قسمت email address قرار میدهیم و از قسمت permission عبارت owner را انتخاب میکنیم. ۴. روی گزینه add کلیک میکنیم. API گوگـل بـه ابـزار گوگـل سـرچ کنسـول سـایت مـا متصـل شـده اسـت. تنهـا مـورد باقیمانده این است که این ویژگی را روی وبسایت خود نیز فعال کنیم.

فعالسازی indexing api در وردپرس

در این بخش به شما آموزش میدهیم چطور indexing api را روی وردپرس راهاندازی کنیم. اگر از سیستم مدیریت محتوای اختصاصی استفاده میکنیم باید با برنامهنویس خود درمورد نحوه فعالسازی api روی سایت صحبت کنیم.

۱. یکی از افزونههای indexing api را از مخزن وردپرس دانلود کنید. (مثلا رنک مث) ۲. از تب افزونه Rankmath گزینه instant indexing را میزنیم و تب Google API Settings را انتخاب میکنیم.

۳. روی گزینه choose files کلیک میکنیم و فایلی که در کامپیوترمان دانلود شـد (فایل json) را در این بخش آپلود میکنیم.

۴. روی ذخیره تغییرات کلیک میکنیم.

در حال حاضر، indexing api گوگل بهصورت کامل برای سایت ما فعال شده است. به همین سادگی، از این بـه بعـد <mark>هـر زمـان کـه تغییری در صفحات وبسایت خود ایجاد</mark> کنیم، گوگل سریعاً از آن مطلع خواهد شد</mark> و از این راه بودجه خـزش و سـرعت اینـدکس افزایش خواهد داشت.

^{بخش ۵} کاهش تعداد آدرسهای سایت

کاهش تعداد آدرسهای غیرضروری سایت، کمک فراوانی به افزایش بودجه خزش میکند؛ **بهویژه اگر سایتی قدیمی و با لینکهای اضافی فراوانی داریم**. حذف لینکهای غیرضروری و کاهش تعداد آن، باعث میشود گوگل تمرکز بیشتری روی لینکهای مهم داشته باشد و نرخ خزش سایت نیز بالاتر برود. برای انجام این کار باید از ابزارهای خزنده استفاده کنیم.

محبوبترین ابزارهای خزنده این ۲ هستند که یک آموزش ویژه نیز برای مهمترین و بهترین ابزار خزنده یعنی اسکریمینگ فراگ نیز آماده کردهایم:

- Screaming Frog .1
 - Ryte ۲

لینک وبسایت خود را در یکی از ابزارها وارد میکنیم. این ابزارها وبسایت ما را کراول میکنند و اطلاعات مفیدی را در اختیارمان قرار میدهند. بخش لینکهای سایت خود را در این ابزار بررسی میکنیم و پس از یافتن آدرسهای اضافی و غیرمفید، برای حذف آنها اقدام میکنیم.

فرض کنیم در وبسایت ما ۲ یا ۳ صفحه با سرعت بسیار پایین وجود دارد. فارغ از دلیل این مسئله، وجود این صفحات کمسرعت و کند موجب کاهش بودجه خزش میشود! بنابراین باید صفحات مختلف، بهویژه صفحات مهم خود را از نظر سرعت بررسی کنیم.



استفاده از WebPageTest

برای بررسی سرعت صفحات مختلف سایت به کمک این ابزار باید مراحل زیر را طی کنیم:

۱. وارد وبسایت start a ... test گزینه Core Web Vitals را انتخاب میکنیم
۲. از قسمت start a ... test گزینه Core Web Vitals را انتخاب میکنیم
۳. آدرس صفحه موردنظر خود را وارد میکنیم
۴. منتظر بررسی صفحه توسط این ابزار میمانیم
۵. نتیجه را بررسی میکنیم.
۵. نتیجه را بررسی میکنیم.
۲. ما درمورد سرعت و کیفیت آن صفحه اطلاعات لازم را میده.
۲. تمامی المنتهای سایت، سرعت لود آنها و نکات تکمیلی در صفحه قرار داده شده
۱۰ است. حتماً اقدامات لازم برای افزایش سرعت آن صفحه و صفحات دیگر را انجام میدهیم تا شاهد افزایش نرخ خزش خود باشیم. برای ایجاد اصلاحات لازم، میتوانیم
۲. با یک برنامهنویس یا متخصص مشورت کنیم.

استفاده از PageSpeed Insights

یکی از ابزارهای مفید گوگل، ابزار PageSpeed Insights است که برای بررسی سرعت و عملکرد یک وبسایت به ما کمک میکند. برای استفاده از این ابزار مراحل زیر را دنبال میکنیم: ۱. وارد آدرس <u>pagespeed.web.dev</u> میشویم ۲. آدرس صفحه موردنظر را وارد و روی گزینه Analyze کلیک میکنیم ۳. منتظر بررسی صفحه توسط این ابزار میمانیم ۴. نتایج تحلیل سایت به ما نشان داده میشود. کمی پایینتر، میزان امتیاز بخش ۴. نتایج الیل سایت به ما نشان داده میشود. کمی پایینتر، میزان امتیاز بخش زمانی است که طول میکشد تا صفحه موردنظر ما به طور کامل بالا بیاید یا اصطلاحاً لود شود.



۵. در این صفحه، تمامی اصلاحات لازم، بهبودها و نکات تکمیلی قرار داده شده است. در صورت داشتن کمی تخصص، میتوانیم بـه کمک ایـن اطلاعـات، عملکرد سـایت را بهینـه کنـیم. اگـر تخصصی نداریـد میتوانـد بـا شـرکت در دوره سـئو تکنیکـال عـلاوه بـر یـادگیری بهینهسـازی سـرعت، مفـاهیم بسـیار پیشـرفته و حرفـهایتری را نیـز آمـوزش ببینید.

بخش V اهل تحقیقات محلی هستید؟!

دوستم تعریف میکرد چند وقت پیش که برای خواستگاری به خانه عروس خانم رفته بودند، پدر عروس لپ تاپش را آورده بود و به دوستم گفته بود: «پسر جون! **من اهل تحقیقات محلی و اینا نیستم!** یوزر و پسورد اینستاگرامت رو بزن بابا جان!!!»

شما هم اگر اهل اینطور تحقیقات محلی نیستید و زمان و انرژی لازم برای بررسی تک تک صفحات سایتتان را ندارید (حتی اگر سایتی کوچک با ۱۰۰ صفحه باشد باز هم کار دشواری است) میتوانید از ابزار همه کاره اسکریمینگ فراگ استفاده کنید. کافی است در تنظیمات این ابزار، آن را به Api خود متصل کنید، تا به صورت یکجا گزارش سرعت تمام صفحات خود را مشاهده کنید.

حذف صفحات بىارزش

معمولاً در تمام سایتها میتوانیم صفحات بیارزشی را پیدا کنیم که به هیچ دردی نمیخوردند؛ این صفحات تنها بودجه خزش سایت را هدر میدهند. با حذف اینگونه صفحات، میتوانیم نرخ بودجه سایت را افزایش دهیم. برای یافتن اینگونه صفحات نیز میتوانیم از ابزارهای خزنده (باز هم اسکریمینگ فراگ!!!) استفاده کنیم و پس از شناسایی، آنها را حذف کنیم.



حذف صفحات اضافي از نتايج جستجو

اگر گوگل صفحه مد نظر ما را ایندکس کرده است باید آن را از بخش نتایج نیز حذف

<mark>کنیم</mark>. برای این کار کافی است وارد گوگل سرچ کنسول شویم و مراحل زیر را طی کنیم:

- Search Console > .1
 - Removals > ۲
 - New Request > .٣
- ۴. 🛛 < واردکردن لینک موردنظر
- Remove this url only > .۵
 - Next > .۶

خش ۸ **ویرایش نقشه سایت**

نقشه سایت، فایلی بهصورت متن است که رباتهای گوگل با خواندن آن میتوانند صفحات و بخشهای مختلف سایت را شناسایی کنند. **اگر در وبسایت ما محتوا و لینکهای زیادی دارد، وجود یک نقشه سایت اصولی و دقیق از واجبات کار است**. مجموع لینکها و صفحات مختلف سایت در این فایل قرار دارد. حال ممکن است بعد از مدتی، ما صفحات زیادی از وبسایت را حذف کرده و یا ریدایرکتهای زیادی انجام داده باشیم؛ اما احتمال این وجود دارد که این آدرسها هنوز در توضیحات نقشه سایت وجود داشته باشند و در نتیجه گوگل و رباتهای خزنده این لینکها را میخزند و با ارور ۴۰۴ مواجهه میشوند یا با چرخیدن در ریدایرکتها، بودجه خزش را هدر میدهند.



یعنی بهطورکلی، بودجه خزش ما صرف خزیدن در صفحاتی میشود که یا وجود خارجی ندارند و ۴۰۴ هستند، یا ریـدایرکت شـدهاند. <mark>بـرای افـزایش سـرعت اینـدکس و بودجـه</mark> <mark>خزش، باید این دسته لینکها و صفحات را از فایل نقشه سـایت حـذف کنیم</mark>. در ادامه به شـما یک افزونـه وردپرس را معرفی میکنیم و بـه شما نشـان میدهیم چطور از آن طریق میتوانیم صفحات بیارزش را از نقشه سایت خود حذف کنیم. نکته: اگر هنوز نقشه سایت خود را در گوگل سرچ کنسول ثبت نکردهاید، حتماً این کار را انجـام دهیـد. ثبت نقشـه سـایت بـه رباتهـای گوگـل کمـک میکنـد تـا بـا سـاختار و اسکلت کلی سایت ما بیشتر آشنا شوند.

ویرایش نقشه سایت با افزونه AIOSEO

ما با کمک افزونه AIOSEO این امکان را داریم که صفحه یا لینک موردنظر خود را از نقشه سایت مخفی کنیم. به این صورت موتورهای جستجو دیگر سراغ آنها نمیروند. کافی است پس از نصب افزونه، مسیر زیر را پیش بگیریم: ۱. وارد قسمت Sitemaps افزونه شویم. ۲. گزینه Advanced Settings را فعال کنیم.

۳. از قسمت Exclude Posts / Pages عبارت موردنظر خود را جستجو میکنیم تا صفحات یا پستهای موردنظر ما لیست شوند.

۴. با انتخاب صفحات و یا پستهای موردنظر، میتوانیم آنها را مخفی کنیم.

نکته: افزونههای دیگری نیز مانند یواست و رنک مث و... نیز وجود دارند.

بخش ۹ این المان المان المان المان المان (المرابض Update Frequency)

میزان تولید محتوا و طراحی صفحات جدید ارزشمند در وبسایت ما با نرخ بودجه خزشمان ارتباط مستقیم دارد. <mark>گوگل وبسایتهایی که محتوای بیشتری را منتشر</mark> میکنند و بهروزتر هستند را با بودجه خزش بیشتری کراول میکند؛

بنابراین یکی از مؤثرترین ترفندها این است که به طور مداوم محتواهای ارزشمندی را برای مخاطبین خود منتشر کنیم و سعی کنیم همیشه وبسایت خود را بهروز نگه داریم.

🕛 محبوبیت سایتهای مرتب و اتو کشیده

نظمدهی به ساختار سایت با کمک لینکسازی داخلی امکانپذیر است<mark>. وجود یک</mark> سـاختار مـنظم و یـک الگـوی مناسـب، بـه خزنـدهها کمـک میکنـد تـا تمـام کوچهپسکوچههای سایت ما را بشناسند و صفحات ما را با سرعت بیشتری کراول کنند و به زبان دیگر، وبسایت و صفحات ما را بهتر درک کنند.

لینکسازی داخلی منظم و هدفمند

برای این کار باید درک درستی از استراتژیهای ساختار سایت داشته باشیم و سعی کنیم صفحات خود را طبقهبندی کنیم و لندینگهای متنوعی بسازیم. <mark>ساخت یک</mark> ساختار به شکل مادر – فرزند میتواند در بسیاری از شرایط برای ما مناسب باشد. بهنوعی که صفحات فرزندها به صفحه مادر لینک میدهند و صفحه مادر نیز به صفحات مهم سایت و صفحات فرزند لینک میدهد. اجراییکردن استراتژیهایی از این قبیل ، به افزایش بودجه خزش کمک میکند. ۱. صفحات مادر را طراحی کنید / مقالههای مادر را بنویسید ۲. صفحات فرزند را طراحی کنید / مقالههای مادر را بنویسید ۴. از صفحات و مقالات فرزند به یکدیگر و صفحه مادر لینک دهید ۴. از صفحه مادر به صفحات مادر دیگر و صفحه مادر لینک دهید ۵. این فرایند را در طول زمان روی سایت پیادهسازی کنید



ایـن یکی از اسـتراتژیهایی اسـت کـه شـما میتوانیـد آن را پیـاده سـازی کنیـد، همـین استراتژی هم نکات بسیاری دارد که در این مطلب نمیگنجد، اگر به اینطور تحلیلهـا و اسـتراتژی چیـدنها علاقهمنـد هسـتید، میتوانیـد **در دوره اسـتراتژی سـئو و لینکسـازی شرکت کنید**.

^{بخش ۱۱} **فشردهسازی یا تغییر کدهای جاوا اسکریپت سایت**

اگر کدنویسی وبسایت ما بهینه نیست (مثلاً فشرده نیست)، نرخ کراول سایت ما بابت همین مسئله در حال هدر رفتن است و سرعت خزش به همین علت کاهش داشته است.

فشردهسازی کدهای css

یکی از راهکارها این است که کدهای css استفاده نشده را شناسایی و آنها را حذف یا جزیره بندی کنیم. انجام این کار علاوه بر افزایش سرعت لود صفحات، نرخ خزش سایت را نیز بالا میبرد و از جنبههای مختلف به رشد وبسایت ما کمک میکند.

شناسایی کدهای css فشرده نشده بهوسیله inspect مرورگر

از سریعترین راهها برای شناسایی کدهایcss استفاده نشده، استفاده از ابزار inspect مرورگر است. برای شناسایی کدهای css بیاستفاده، مراحل زیر را پیش ببرید: ۱. ورود به صفحه موردنظر ۲. کلیک راست روی یک بخش از صفحه ۳. انتخاب گزینه inspect ۹. کلیک روی گزینه سه نقطه موجود در پایین صفحه (در کنار گزینه console ۵. انتخاب گزینه Coverage پس از انجام این مراحل، لیست کدهای css سایت بر اساس درصد استفاده به شما نشان داده میشود که با توجه به شرایط، آنها را حذف یا جزیرهبندی میکنیم.

بخش ۱۲ دلبری کنید و محبوب باشید

سایتها و آدرسهای مهمتر و محبوبتر، نرخ خزش بالاتری نیز دارند</mark>. دلیل این مسئله میزان اعتبار و محبوبیت زیادی است که این آدرسها توانستهاند در طی زمان کسب کننـد. در ادامـه سـه راه را بـرای افـزایش محبوبیـت و اعتبـار صـفحات بررسـی میکنیم. ۱. لینکسازی خارجی ۲. استفاده از سوشال سیگنالها

۳. افزایش تعامل کاربران

لينكسازي خارجي

هرچقدر وبسایتهای مرتبط و محبوب، صفحات سایت ما را پیشنهاد دهند (به آنها لینک دهند) اعتبار و محبوبیت ما بیشتر خواهد شد. پیشنهاد میکنیم ریپورتاژ آگهی را بهصورت اصولی یاد بگیرید و با انجام آن، اعتبار و ارزش دامنه و سایت خود را نزد گوگل افزایش دهید. در دوره لینکسازی، به صور مفصل نکات مربوط به انتخاب رپورتاژ آگهی توضیح داده شده است.

استفاده از سوشال سیگنالها

قراردادن لینک یک مطلب یا صفحه جدید در شبکههای اجتماعی مختلف از جمله توییتر، در میزان سرعت ایندکس و افزایش نرخ خزش سایت تأثیرگذار است. با این کار از شبکههای اجتماعی کاربران زیادی را به سایت وارد میکنیم و همین نشان از محبوبیت و اهمیت آن صفحه است. **اگر مهم و محبوب نبود چرا باید از شبکههای اجتماعی ورودی بگیرند؟** گوگل از این موضوع کاملاً مطلع است و به همین دلیل هم اگر این کار را به درستی انجام دهیم، بودجه خزش را افزایش میدهیم.

پیشنهاد میکنم برای استفاده از ایدههای مؤثر، حتماً دو فصل «لینک سازی رایگان» را در دوره لینک سازی آکادمی وبسیما مشاهده کنید.

بخش ١٣ لطفا يک کاری کنيد!

هرچقدر کاربران با صفحه ما تعامل بیشتری داشته باشند، یک سیگنال قوی برای نشان دادن محبوبیت، ارزشمندی و کاربردی بودن صفحه ماست<mark>.</mark> تعامل میتواند به هر روشی انجام شود:

- کاربر وارد یکی از لینکهای صفحه ما شود
- پست ما را لایک کند یا برای ما نظر بگذارد
 - فرمی را پر کند و ...

بنابراین، **باید با استفاده تکنیکها و ترفندهای جذاب و با توجه به شرایط، تعامل** مخاطب را با محتوایمان افزایش دهیم.

مش ۱۴ تغییر سرعت ایندکس بهصورت دستی

یکی از کارهایی که با انجام آن احتمال دارد سرعت ایندکس صفحات سایت ما را افزایش یابد و موجب افزایش بودجه خزش شود، تغییر سرعت ایندکس از طریق سرچ کنسول است. سایتهایی که مطالب زیادی را در طول روز در وبسایت خود منتشر میکنند از این گزینه استفاده میکنند. سایتهای خبری از این دسته هستند. **این گزینه برای سرچ کنسول جدید قرار نگرفته است**؛ اما در ادامه به شما آموزش میدهیم چطور به آن دسترسی پیدا کنید. برای تغییردادن سرعت ایندکس از طریق

سرچ کنسول، مطابق گامهای زیر پیش میرویم:



https://support.google.com/webmasters/answer/۴۸۶۲۰?hl=en . آدرس (ا باز میکنیم.

۲. در توضیحات تیتر ««Limit the crawl rate، روی عبارت Crawl Rate Settings . pageکلیک میکنیم و اگر وارد اکانت گوگل خود نشدهایم، وارد میشویم.

۳. در این بخش دو گزینه را مشاهده میکنیم. روی Limit Google's maximum crawl rate کلیک میکنیم.

۴. در این بخش، محور سرعت ایندکس باز میشود. ما میتوانیم درجه کراول یا خزش سایت توسط رباتهای گوگل را از این قسمت تنظیم کنیم. پس از تغییر سرعت کراول به میزان دلخواه، روی saveکلیک میکنیم تا تنظیمات ذخیره شوند.

نکته: **افزایش نرخ خزش از این قسمت لزوماً باعث افزایش آن نمیشود** و به تمام آن فاکتورهایی که تا اینجا با آنها آشنا شدید بستگی دارد. اما اگر میزان خزش در سایت را کمتر کنیم، شانس بیشتری داریم تا گوگل حرف ما را قبول کند.

^{بخش ۱۵} آستینها را بالا بزنید

ما میتوانیم آدرس صفحه مورد نظر را بهصورت دستی در گوگل سرچ کنسول ثبت کنیم. برای انجام این کار پس از ورود به ابزار گوگل سرچ کنسول، وارد بخش URL Inspection میشویم و آدرس صفحه مورد نظرمان را وارد و سوس روی گزینه Request Indexing کلیک میکنیم تا رباتهای گوگل صفحه را بخزند و آن را ایندکس کنند.

تا اینجای کار تمامی راهکارهای اصلی و مهم افزایش سرعت ایندکس و بودجه خزش را بـه صـورت کامـل و جزئـی بررسـی کـردیم؛ <mark>بـروز نگهداشـتن سـایت، بهینهسـازی نسـخه</mark> موبایـل و اسـتفاده از API گوگـل، مهمتـرین نکاتی بـود کـه بایـد بـه آنهـا توجـه ویـژهای داشـته باشیم. حـالا قصـد داریـم در فصـل بعـدی، تلـههای خطرناکی را بـه شـما معرفی کنیم که هر کسی ممکن است گرفتار آنها شود و سایت خود را نابود کند.



🖉 بررسی تلههای خزنده در سایت و رهایی از آسیب آنها



حواسمان را باید خوب خوب جمع کنیم؛ ما فقط یک متخصص سئو نیستیم و باید مو را از ماست بیرون بکشیم؛ **ما کارآگاهی ریزبین هستیم!** دنیایی که در آن فعالیت میکنیم پر است از خطرهای جـدی، تلـههای کشـنده، اشـتباههای مرگبـار! اگـر در ایـن مسیر یا کج بگذاریم، کارمان تمام است ...

خب خواستم کمی با شما شوخی کرده باشم و موضوع این فصل را جنایی کنم! نگران نباشید اگر ایراد و مشکلی در سایت شما وجود داشته باشد یا اگر راهی را اشتباه رفته باشید و مشکلی را در سایت خود به وجود آوردهاید، هیچ عیبی ندارد؛ عیب آنجاست که مشکلات را نشناسیم و راه حل آنها را بلد نباشیم.



تارهایی برای شکار عنکبوت!

در سایت ممکن است چندین تلـه خطرنـاک بـرای شـکار کـردن رباتهـای گوگـل وجـود داشته باشند، این تلهها، رباتهای خزنده گوگل را شکار میکنند و بودجه خزش سایت به کلی نابود میشود. قصد داریم این تلهها را به شما معرفی کنیم و راه از بین بردن هرکدام را به شما آموزش دهیم. باید در سایتمان به دنبال این تلهها بگردیم و هـر کجـا که آنها را دیدیم فوراً نابودشان کنیم!

تلههای شکارچی کجا هستند؟

اولین قدم برای نابودی تلهها یافتن مکان آنهاست، مهمترین فضاهایی که این تلهها

- حضور دارند شامل این صفحات هستند:
 - ۱. صفحه ورود کاربران
 - ۲. صفحه ثبتنام
 - ۳. صفحه علاقهمندیهای کاربر
 - ۴. صفحه پرداخت نهایی
 - ۵. سبد خرید
 - ۶. صفحه ادمین
 - ۷. نتایج جستجو در سایت
 - ۸. نتایج مرتبسازی (Sort) و ...

نکته: البته تلهها را در فضاهای بیشتری نیز میتوانیم پیدا کنیم. در ادامه به بررسی موردی هر کدام میپردازیم:

بخش کردن مینهای انفجاری:

ما باید با احتیاط کامل، در سایت خود حرکت کنیم و به دنبال این تلههای انفجاری بگردیم و آنها را خنثی کنیم: ۱. تلههای زمانی: (تقویم و رویدادها) ۲. جهتیابی چند وجهی (ایراد در navigation سایت) ۳. ایرادات مربوط به ریدایرکت سایت ۴. ایجاد صفحات زیاد با جستجو داخلی ۵. توجه به پروتکل های امنیتی سایت ۶. ایندکس آدرس دارای ویژگی ۷. وجود سشن آیدی در آدرسها ۸. پلاگینهای فشردهسازی عکس

گمشده در زمان (تقویم و رویدادها)

میخواهیم یک کار خفن انجام دهیم! قصد داریم مدت زمان پایان تخفیف دوره هیولای سئو را با استفاده از یک رویداد شمار به کاربران سایتمان اعلام کنیم؛ این رویداد شمار، ثانیه به ثانیه در حال تغییر است و مدت زمان باقی مانده تخفیف را نشان میدهد (مانند ثانیه شمار یک بمب ساعتی)، **حالا محتوای این صفحه، در لحظه تغییر میکند** باعث میشود که رباتها نرخ خزش بیشتری را صرف آن صفحه کنند. اگر در صفحات سایت خود، تقویم و یا یک رویداد تاریخدار دارید، برای جلوگیری از ایجاد چالشهایی جدی حتماً به این مسئله توجه داشته باشید.

غیرفعال کردن ایندکس بخش تایمر

برای رفع این مشکل باید ایندکس بخش تایمر صفحه مورد نظر را غیرفعال کنید. یعنی باید قسمت تایمر صفحه را noindex کنید تا از بروز چنین مشکلی جلوگیری شود.

ردپای مشکوک (ایراد در navigation سایت)

کمی به مسیر زیر دقت کنید: آیا میتوانید تشخیص دهید که صفحه انتهایی به چه محتوایی میرسد؟ آیا آن را درک میکنید؟

دیجیکالا / اسباب بازی، کودک و نوزاد / سرگرمی و آموزشی / اسباب بازی / جغجغه، عروسک و مدل / عروسک بـه ایـن فسـمت از سـایتها «جهتیابی» یا «INavigationi» می دوینـد. ایـن فسـمت نقشهای را که مخاطب دنبال کرده تا به صفحه مد نظر خود برسد را مشخص میکند، حالا اگر در ایـن بخش از سـایت، بعـد از «اسـباب بـازی» لینک صفحه «لـوازم مصرفی خـودرو» وجـود داشت، ایـن مسـئله کمـی مشـکوک نمیشـد؟ آیـا مـا گـیج نمیشـدیم؟ میتوانستیم درک درستی از این سایت داشته باشیم؟

وجود مشکل در بردکرامب (Navigation) سایت باعث میشود که خزندهها گیج شوند و مسیریابیها را به صورتی اشتباه انجام دهند. از همه مهمتر اینکه آنها سایت ما را به درستی درک نمیکنند. علاوه بر اینکه **ساختار منظم و هدفمند ما از بین میرود و شانس کسب رتبههای برتر را از دست میدهیم**، در یک تله خزنده نیز گیر افتادهایم و بودجه خزشمان نیز هدر رفته است.

پاک کردن مسیریابی اشتباه

برای برطرف کردن این مشکل، مسیریابی اشتباه را به طور کلی پاک کنید. اگر به مشکلی برخوردید که نتوانستید مسیرهای اشتباه را حذف کنید، مثلاً دسترسی نداشتید یا تعداد آنها خیلی زیاد بود، بهطور کلی برد کرامپ را حذف کنید و از یک متخصص برای ساخت یک برد کرامپ جدید و اصولی کمک بگیرید. در غیر این صورت از راهحل زیر استفاده کنید.

استفاده از TEXT FRAGMENT

اگر تعداد مسیرهای اشتباه سایت زیاد است و حذف تک تک آنها بهصورت دستی امکانپذیر نیست، بهجای حذف آنها، میتوانیم مسیر اصلی را به خزندهها معرفی کنیم. برای انجام این کار باید از ویژگی «text fragment» استفاده کنیم. به کمک آن میتوانیم مسیر صحیح و مورد نظر خود را به صفحه اضافه کنیم. در این صورت خزندهها با خواندن آن، مسیر درست سایت را درک میکنند.

هزارتوهای پیچ در پیچ (ریدایرکتهای اشتباه)

معمولاً یکی از مشکلات رایج در سایت مربوط به ایرادهای ریدایرکتهای آن است، در ادامه رایجترین و مهمترین ایرادهای مربوط به ریدایرکتها را بررسی میکنیم:



زنجیرہ بیپایان ریدایرکت

زنجیره بیپایان، به ریدایرکتهایی گفته میشود که هرگز تمامی ندارند! احتمالاً پیش آمده باشد که وقتی وارد صفحهای شدهاید این ارور به شما نشان داده شده است: «...this page isn't working! example.com redirected you too many times» این خطا نشان دهنده وجود تعداد زیادی ریدایرکت است، مثلاً ۵ صفحه را در نظر بگیرید که همگی روی یکدیگر ریدایرکت شده باشند. اگر این زنجیرههای ریدایرکت را پیدا نکنیم و مشکل آن حل نشود با چالشهایی جدی روبرو خواهیم بود. (راهحل این مشکل ساده و آسان است: آنها را حذف میکنیم!)

نکته: رباتهای گوگل، زنجیرههای ریدایرکت را تشخیص میدهند و به همین علت بعد از چند بار مشاهده این خطا، دیگر آن صفحه را کراول نمیکنند. البته این نکته را در نظر بگیرید که وجود چندین زنجیره ریدایرکت نشان دهنده کیفیت بد سایت است و اعتبار سایت را کم میکند.

طولانىبودن حلقه ريدايركت

حلقه ریـدایرکت زمـانی رخ میدهـد کـه مـا آدرس ۱ را بـه آدرس ۲ ریـدایرکت مـیکنیم؛ درصورتیکه خـود آدرس ۲ نیـز قـبلاً بـه یـک آدرس دیگـر ریـدایرکت شـده است! بـه ایـن اتفاق، ایجاد یک زنجیره ریدایرکت میگویند؛ امـا **اگـر در پایـان url هـدفی وجود نداشته باشـد شـما درگیـر یـک تلـه خزنـده شـدهاید**. داشـتن حلقـه ریـدایرکت اضـافه بـهطورکلی میتواند نـرخ کـراول شـما را هـدر دهـد و خـزش در صفحات سـایت را دچـار مشکل کنـد. برای حل این مسئله باید ریدایرکتهای اضافی را در مسیر حذف کنیم.



بهروزرسانی ریدایرکتها از http به http

وجـود ریـدایرکتهای اشـتباه و یـا بـروز نشـدن ریـدایرکتها از http بـه http **موجـب ساخت یک لوپ یا حلقه ریدایرکت میشود**. حتماً این ریدایرکتها را بررسی کنید.

باز شدن سایت با www یا بدون آن

وبسایت خود را یکبار با www و یکبار بدون آن بررسی کنید:

- www.websima.academy .1
 - websima.academy .Y

سایت باید فقط با یکی از این دو آدرس باز شود<mark>، در غیر این صورت گرفتار تله خزنده</mark> شدهایم و باید یکی از آنها را روی دیگری ریدایرکت کنیم.

استفاده از ابزار تشخیص حلقه ریدایرکت

یکی از ابزارهایی که به شما کمک میکند تا حلقههای ریدایرکت را شناسایی و آنها را برطرف کنید، نرمافزار Screaming Frog است. کافی است این ابزار را نصب کنید و پس از واردکردن آدرس سایت موردنظر خود، برای بررسی خطای حلقه ریدایرکت مراحل زیر را پیش بگیرید:

Reports > Redirects > Redirect Chains

فلفل نبین چه ریزه (ابزارک جستجو در سایت)

یکی دیگر از بخش های مهمی که باید بررسی کنیم، ابزارک جستجوی داخلی سایت است. وقتی در این بخش عبارتی را جستجو میکنیم، تمام صفحات مختلف سایت و مرتبط با آن، لیست خواهد شد. این یک قابلیت بسیار عالی برای کاربران ماست و باعث میشود آنها به راحتی و با سرعت بالایی به محتوای مورد نظر خود برسند.



اما اگر کمی دقت کنید **این قابلیت باعث ایجاد صفحاتی با لینکهایی منحصر به فرد میشود** و این یعنی ایجاد صفحات بسیار زیاد با محتواهای تکراری!!! بنابراین اگر در وبسایتمان از این قابلیت استفاده میکنیم، باید دسترسی موتورهای جستجو به این بخش را بهطورکلی ببندیم.

بستن دسترسی موتورهای جستجو به کمک فایل robots.txt

همانطور که قبلاً هم به این قسمت اشاره کردیم، هر زمان که قصد داشته باشیم دسترسی موتورهای جستجو را به بخش مشخصی از وبسایت و یا حتی کل سایت ببندیم، میتوانیم از فایـل robots.txt اسـتفاده کنـیم؛ بایـد وارد ایـن فایـل شـویم و بخش مورد نظر را «disallow» کنیم.

نکته: در سیستم مدیریت محتوای وردپرس میتوانیم از افزونههایی برای ساخت و ویرایش فایل دستورالعمل رباتها استفاده کنیم؛ البته پیشنهاد ما این است که این فایل را بهصورت دستی و بدون دخالت افزونه ویرایش کنید. (در دوره آموزش گوگل سرچ کنسول این قسمت را به طوری اختصاصی و حرفهای آموزش دادهایم)

کار نصفه و نیمه ممنوع (پروتکلهای امنیتی)

اگر در سایتمان پروتکل امنیتی (SSL) را فعال کنیم، تمام آدرسهای Http با ریدایرکت به Https تبدیل میشوند. اگر سایتمان لینکها یا صفحات بیکیفیت و بیارزش زیادی داشته باشد، ممکن است این ریدایرکتها به درستی انجام نشوند. مثلاً تمام صفحاتمان به HTTPs تبدیل نشوند. دقت کنید که همین مسئله ساده باعث بروز مشکلاتی جدی برای رباتها میشود و چالشهایی جدی را برایمان رقم میزند. مثلاً HTTP

اگر از این تصویر، در یکی از صفحات سایت خود استفاده کرده باشیم چه اتفاقی رخ خواهد داد؟ در صفحه امن ما، محتوایی نا امن وجود دارد و همین ایراد ساده، مشکلی جدی برای صفحه است. یا ممکن است صفحه خرید عطر کرید اونتوس به جای اینکه جایگزین نسخه Https خود شود، روی صفحه اصلی سایت ریدایرکت شود؛ این مشکلات معمولاً زمانی رخ میدهند که ما سیستم مدیریت محتوا یا سرور سایتمان را تغییر میدهیم. بنابراین **پس از انجام این کار، حتماً آدرسها را بررسی کنید تا از انتقال** درست آدرسها مطمئن شوید.

هر گردویی گرد هست (ایندکس پارامترها)

این مسئله برای فروشگاههای آنلاین و بیشتر در بخش فروشگاهی رخ میدهد. اگر تابهحال وارد یـک فروشـگاه اینترنتـی معتبـر شـده باشـید، در یـک بخـش از صـفحه فیلترهای مختلفی را میبینید که به کمک آنها محصولات را بر اساس مدل، قیمت، رنگ و... محدود میکنید.

اگر قسمت فیلترها را noindex نکنیم یا بهاصطلاح، موتورهای جستجو قابلیت ایندکس این بخش را داشته باشند **تعداد زیادی لینک و صفحه بیکیفیت ایندکس میشود** که تمامی آنها بهعنوان یک تله عنکبوتی عمل میکنند. پارامترها انواع مختلفی دارند مانند:

 فیلتر محصولات: برای مثال در یک فروشگاه اینترنتی میخواهیم صرفاً محصولاتی تا سقف قیمت ۵ میلیون تومان را ببینیم یا میخواهیم محصول موردنظرمان دارای شکل، مدل و حالت بهخصوصی باشد. تمامی فیلترهای انتخابی ما به ادامه url اضافه میشوند. برای مثال:

mrpashmak.com/mobile?brand=samsung&storage=۲۵۶&color=white



در این نمونه ما میخواهیم فقط موبایلهای برند سامسونگ با حافظه ۲۵۶ گیگابایت و سفید رنگ را مشاهده کنیم. حالا بررسی کنیـد چنـد نمونـه از ایـن آدرسهـا را بـا استفاده از فیلترها میتوانیم بسازیم؟ به نظر من که واقعاً تعدادش زیاد است! حال اگر کراولرهـای گوگـل ایـن آدرسهـای منحصـر بـه فـرد و بیشـمار را ببینند و قصـد ایندکس کـردن آن را داشته باشند، چقـدر طول میکشد؟ بنابراین بایـد هرچـه سـریعتر این مشکل را رفع کنیم. اگر دقت کنید دقیقاً قبل از شروع شدن آدرس فیلترها یک علامت سؤال «؟» در لینک

قرار میگیرد. ما دقیقاً با توجه به همین ویژگی مشخص میکنیم که <mark>آدرسهایی که</mark> دارای علامت سؤال هستند، ایندکس نشوند</mark>؛ به همین سادگی!

استفادہ از تگ no index

یکی از راه حلهای این مسئله افزودن تگ noindex به لینکهای دارای پارامتر است.

بستن دسترسی موتورهای جستجو از طریق فایل robots.txt

کافی است وارد فایل robots.txt شوید و دسترسی موتورهای جستجو را به صفحات دارای فیلتر ببندید.

قدیمی اما ممکن (سشن آیدیهای URL)

به زبان ساده، سشن به معنای مدت زمانی است که یک کاربر در وبسایت حضور دارد. تمامی اقدامات و فعالیتهایی که مخاطب در این مدت زمان انجام میدهد در یک آدرس اختصاصی نگهداری میشود و برای هر کاربر متفاوت است که به آن سشن آیدی (Session ID) میگویند.

در حال حاضر بیشتر وبسایتها به جای سشن آیدی از کوکی استفاده میکنند.

همانطور که گفتیم، **آیدی بهصورت یک آدرس منحصربهفرد است که بعد از آدرس سایت قرار میگیرد**. کراولرها و یا خزندهها، هر آیدی که بعد از آدرس ظاهر میشود را بهعنوان یک لینک جدید محاسبه میکنند و این اتفاق باعث ساخت تعداد زیادی لینک بیارزش میشود. هرچقدر تعداد آیپیهای ورودی سایت بیشتر باشد، تعداد این لینکها نیز بیشتر است.

تعریف session ID بهعنوان یک ویژگی url

session ID را بهعنوان یک پارامتر در آدرس تعریف کنید. اگر به یاد داشته باشید، در قسمت ایندکس شدن آدرسهای دارای پارامتر، درمورد این صحبت کردیم که ما باید فیلترها و پارامترها را noindex کنیم تا موتورهای جستجو آنها را کراول نکنند.

اگـر هـر سشـن آیدی را بـهعنوان یـک پـارامتر تعریـف کنـیم، از آنجـا کـه پارامترهـا را noindex کردهایم و یا از طریق فایل robots.txt دسترسی موتورهای جسـتجو را بـه آن بستهایم، دیگر نگران ساختهشدن لینکهای زیاد و بیارزش نخواهیم بود.

تغییر ساختار url از طریق سورس کد

در ساختار برخی از سایتها سشن آیدی بهصورت embed در خود url وجود دارد پس تعریف آنها به عنوان یک پارامتر، راهحل درستی نیست. در ایان شارایط، بایاد ویرایشهای لازم را از طریق سورس کد سایت به وجود بیاوریم. برای نوشتن کدهای لازم با یک برنامهنویس مشورت کنید.

نکته: زمانی که در یک وبسایت، سشن آیدی به صورت embed در url وجود داشته باشد، آن را در قسمت نتایج موتورهای جستجو از جمله گوگل مشاهده میکنیم. برای مثال ممکن است در نتایج گوگل، عبارت jsessionid را پس از url خود صفحه ببینیم.

آمدم ثواب کنم کباب شدم (پلاگین فشردهسازی تصاویر)

افزونههایی وجود دارند که با آنها، عکسهای وبسایت یا صفحههای مختلف سایت را فشرده میکنیم. این افزونهها برای وبسایتهای ساخته شده با cms وردپرس بسیار محبوب هستند و اگر قصد استفاده از آنها را داریم باید خوب حواسمان را جمع کنیم! مشکلاتی که ممکن است این افزونهها ایجاد میکنند:

ساخت ورژنهای مختلف صفحه

برخی از این افزونهها با فشردهسازی فایلهای مختلف یا عکسهای یک صفحه، نسخههای مختلفی از آن عکس را طراحی میکنند. برای مثال عکسهای موجود در یک صفحه را به ابعادی مختلف برای دستگاههای مختلف تبدیل میکند و کراولرها تمامی نسخههای مختلف تولید شده را بررسی میکنند و هرکدام از آنها را بهعنوان یک صفحه مجزا میبینند. اگر احساس میکنیم نرخ کراول زیاد است و بیشتر آن صرف ایندکس عکسهای سایت میشود، باید افزونههای مربوط به فشردهسازی عکس و فایلهای سایت را بررسی کنیم.

^{بخش ۳} کک توی تنبان سایتت افتاد؟

خب شاید پس از اینکه با اینهمه مشکل و تله و بمب انفجاری و تلههای کشنده آشنا شدهاید، نگران سایت خود هستید. برای اینکه مشکلات خزش و ایندکس سایت خود را بررسی کنید و از سلامت سایت مطمئن شوید چندین راه وجود دارد:



امتحانش مجانى است

ابزارهای مختلفی مانند Screaming Frog و Scrapy طراحی و عرضه شدهاند که به کمک آنها از وبسایت خود تست کراول میگیریم و لینکهای مختلف سایت را میبینیم. پس از انجام تست، تمامی لینکها را تاحدامکان بررسی میکنیم و ذرهبین خود را در دست میگیریم تا ببینیم **چه لینکهایی ایندکس میشوند و در صورت وجود هر مورد مشکوکی دلایل آن را پیدا و حل میکنیم**.

نکته: این ابزارها علاوه بر بررسی لینکها و کمک به ما در رفع تلههای خزنده، به بهینهسازی سایت از لحاظ فنی و سئو کمک میکنند.

حرفهای جستجو کنید

در صفحه جستجوی گوگل آدرس سایت خود را با استفاده از عبارت «site» جستجو میکنیم؛ مثلاً (site:websima.academy) تمام صفحات ایندکس شده سایت ما، لیست میشود. به همین راحتی!

یک ترفند دیگر این است که کد «:inurl» را به آن اضافه کنیم و قسمتی را بنویسیم که بـه دنبـال آن هسـتیم؛ مـثلاً بـرای اینکـه متوجـه شـویم سشـن آیدیهـا در سـایت مـا اینـدکس میشـوند یـا نـه، عبـارت «sessionid» را بـه آن اضـافه مـیکنیم. (مـثلاً (site:websima.academy inurl:sessionid) لینکهـایی کـه بـرای مـا لیسـت میشوند دارای سشن آیدی هستند و باید مشکل آنها را حل کنیم.

با استفاده از ایـن کـد هـر چیـزی را در هـر سـایتی میتوانیم جسـتجو کنـیم. مـثلاً مـن میخـواهم در سـایت آکـادمی وبسـیما تمـام آدرسهـایی کـه در آنهـا عبـارت «Crwal» وجود دارد را پیدا کنم: site:websima.academy inurl:crawl





گوگل تمام آدرسهای سایت آکادمی وبسیما را بررسی کرد و آدرسهایی که در آنها عبارت crawl وجود داشت را به من نشان داد؛ چندین عکس و یک صفحه محتوا در واژه نامه سئو را برای من لیست کرد. (سخنی از نویسنده: «باحاله نه؟ به نظرت چه کارهای جذابی باهاش میشه کرد؟؟؟؟»)

خیلی خفن باش (Log File)

گزارشهـای وبسـایت در فایـل لاگ ثبـت میشـود. فایـل لاگ شـامل تمـام درخواستهایی است که از طرف بازدیدکنندگان و کراولرها (خزندهها) بـه سـرور سـایت فرستاده شده است. این فایل را بررسی میکنیم و در صورت وجود آدرسهای اشـتباه، مشکوک و دارای تله عنکبوتی، برای رفع آنها اقدام میکنیم.



آخرین سخن

چقدر لحظاتی خوشی را تا اینجا سپری کردیم! نوشتیم و خواندید و خندیدیم و یاد گرفتیم! حالا وقت آن رسیده است که آستینها را بالا بزنید و به جان سایتهای خود بیفتید... خالق این کتاب همیشه میگوید:«فرق است بین کسی که فقط میداند و کسی که به دانستههایش عمل میکند!» شروع سئو بدون دانش بد است و توقف سئو صرفا به دلیل یادگیری مضر! بهترین راه این است که یاد بگیرید و در همان لحظه با عمل به دانش خود تجربه کسب کنید.

من یک درخواست هم از شما دارم؛ حالا نوبتی هم باشد نوبت نظر دادن است! کدام بخشهای کتاب را دوست داشتید؟ چه قسمتهایی برایتان مفید بود؟ چه نقاط قوت و ضعفی را حس کردید؟ و هر نظر دیگری که دارید را در <mark>صفحه کتاب زمستان سرد ایندکس</mark> <mark>در سایت آکادمی وبسیما</mark> ثبت کنید. ما تمام نظرات را به دقت بررسی میکنیم و به سوالات شما پاسخ میدهیم. قطعا نظرات شما برایمان ارزشمند است و در مسیر تکامل و پیشرفت از آنها استفاده خواهیم کرد.

ما در شبکههای اجتماعی #زمستان_سرد_اینـدکس را راه انداختـهایم و محتواهـای آن را دنبال میکنیم. اگر بخشی از کتاب را پست کردیـد یا نظـری درباره آن دادهایـد، خوشحال میشویم از این هشتگ استفاده و ما را منشن کنید تا از آن با خبر شویم.

www.linkedin.com/in/abolfazlparsaie/

www.linkedin.com/company/websima-academy/