The Beginner's Guide to Googlebot Optimization

Executive Summary

Among the top tactics to improve your brand website's organic rankings sits Googlebot optimization, a way to play well with Google's web crawler in order to improve your drug or disease-state website's results. This POV defines the role of Googlebot in SEO and offers a checklist of how to improve your results. (Don't worry, Googlebot is not a bad bot and doesn't count as non-human traffic. Googlebot is a good bot that you want crawling your site.)

Background

Many folks know about search engine optimization, which is essential for a website to rank well in search engines. But chances are, you may not have heard of Googlebot optimization, which focuses on how Google's crawlers access your site. Technical SEO is a very important and often overlooked piece of the SEO process. In order for a site to rank well, all the technical aspects need to be working correctly so Google is receiving the best possible signals from your website.

Optimizing for Googlebot is an essential step in the technical SEO process and a must if you want to rank well in today's search landscape.

To get organic traffic, a website must be in Google's index. And a site cannot get into the index if it cannot be crawled. Thus a site's crawlability is the crucial first step to ensuring its searchability within the index.

What is Googlebot?

Googlebot is Google's search bot, also known as a spider or crawler, which crawls the web and creates an index. Googlebot will crawl every page it's allowed access to, and adds that page to the index where it can be returned by a user's search query.

Understanding how Google crawls your site is crucial to understanding Googlebot optimization. Here are the basics:

1. **Googlebot spends more time crawling sites with higher PageRank**

> " Don't worry – Googlebot is not a bad bot."

PageRank (PR) is a quality metric invented by Google's owners Larry Page and Sergey Brin. PageRank is a 0 to 10 scale that assigns a number to a web page based on that page's importance, reliability and authority on the web according to Google.

To put PageRank into perspective, Google.com is a 9, most consumer brand pages are around 3, most pharma company brand pages are 2 or 3, and health related websites like WebMD or Healthline are 7s. The breadth of content, the amount of inbound links and website history are all factors that go into PageRank. The amount of time that Googlebot gives to your site during the crawling process is called "crawl budget". The greater a page's authority or PageRank, the more crawl budget it receives.

2. Googlebot is always crawling your site

Fresh, consistent content always gains Googlebot's attention and improves the likelihood of ranking better. New backlinks and social mentions will do the same. It's important to note that Googlebot does not crawl every page on your site all the time.

3. Googlebot first accesses a site's robots.txt file to find out the site's crawling rules

Your robots.txt file tells Googlebot where it cannot go. Pages with sensitive user information and checkout pages are common in a robots file. Any pages that you would not want in the index would get added to your robots file. Pages that are within the robots file will not be crawled or indexed by Googlebot.

4. Googlebot uses the sitemap.xml to discover any and all areas of the site to be crawled and indexed

On the flip side, Googlebot uses your XML sitemap to find a list of pages that you do want indexed. Due to the many ways in how sites are structured, the crawler might not get to every page or section. In this case, Googlebot will refer to your sitemap to ensure it gets to all pages and crawls them.

5 Steps for Googlebot Optimization

Googlebot optimization comes a step before SEO when it comes to optimizing a website. Below are five steps you should take to ensure your site is ready for Googlebot to crawl.

1. Don't use fancy coding languages

Googlebot doesn't crawl JavaScript, iframes, DHTML, Flash, and Ajax content as well as HTML. Stick to the basics (HTML).

2. Enhance your robots.txt file

Your robots.txt file is essential because it serves as a guide for Googlebot. Googlebot will spend its crawl budget on any pages on your site. You need to tell Googlebot where it should and shouldn't spend that budget. So if there are any pages or sections of your site that you do not want indexed, be sure to add them to your robots.txt file. The less time Googlebot spends on unnecessary pages of your site, the more it can crawl more important pages of your site.

> "Googlebot optimization comes a step before SEO when optimizing a website."
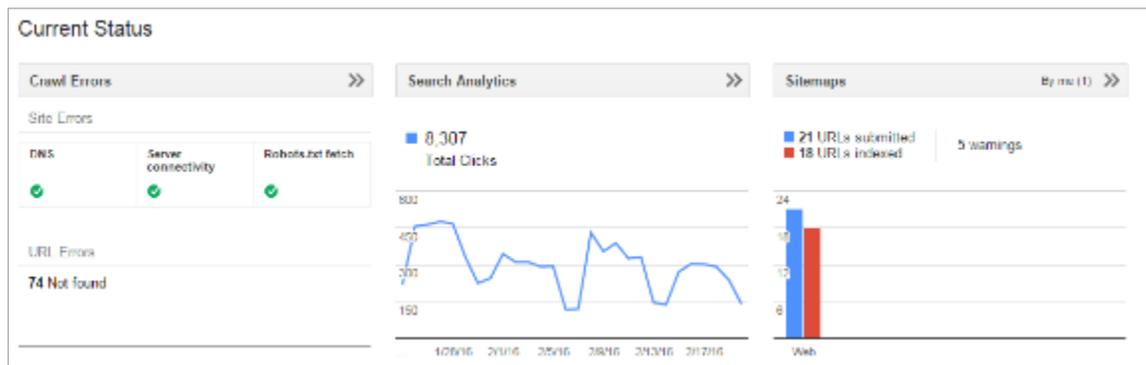
3. Create fresh content

Content that is crawled more often is likely to gain more traffic. While PageRank is the main factor in crawl frequency, studies show it's likely that PageRank becomes less important when compared with fresh content.

Another important piece of Googlebot optimization is to get your low ranked pages crawled as often as possible. One way to do this is to update these articles often as new information or data becomes available.

4. Use internal linking

Internal linking is, in essence, a map for Googlebot to follow as it crawls your site. The bot will follow your internal links until it hits a dead end, at which point it will refer to your sitemap. The more tight-knit your internal linking structure, the better Googlebot will crawl your site.

5. Create an XML sitemap

Your sitemap is one of the clearest messages to the Googlebot about how to access your site. A sitemap does exactly what the name suggests — serves as a map to your site for the Googlebot to follow.

Sitemaps ensure that all the areas of your site that you want indexed will get crawled.

Analyzing Googlebot's performance on your site

The best part about Googlebot optimization is that you don't have to guess to see how your site is performing with the crawler. By implementing Google Search Console (formerly Webmaster Tools), we can see exactly how Googlebot is indexing the site.

> "Content that is crawled more often is likely to gain more traffic. "

The above screenshot is the dashboard you'll see upon logging in. It provides a high-level view of the technical pieces, including any crawl errors, organic search impressions, as well as indexed URLs.

Search Console also allows us to see potential errors and warnings that we should address. In this screenshot, we are in the XML sitemaps section and can see when Google last processed the sitemap and how many URLs are indexed.
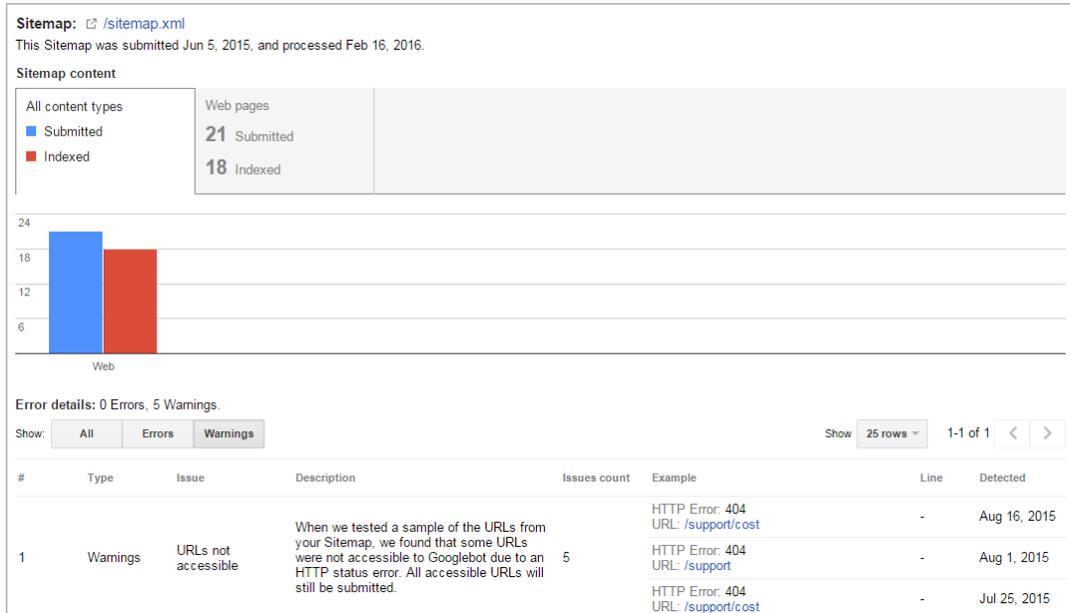
Conclusion

If you want to improve your site's performance and organic visibility, Googlebot optimization is a great place to start.

By considering how Googlebot crawls a site during the early stages of development, you ensure that all of the great content and design work produced for the site is able to be indexed and ranked in results.

To be indexed and returned in search engine results, a site must get crawled - so give the crawlers the directions and map they need to efficiently crawl and index your website.

Analysts:

Brian Cox
SEO Supervisor
CMI Media LLC