



The Technical SEO Guide

By Charles Floate

Table of Contents -

Technical SEO In A Nutshell	3
What's The Difference Between Technical SEO And OnPage?	3
Why Is Technical SEO Important?	4
So How Does Technical SEO Work?	4
Search Spiders	5
Crawling	5
Crawl Budget	5
Log File Analysis	6
Load Balancing	8
Indexing	8
Tools We'll Need	9
Factors to Analyze in Technical SEO	10
Site Structure and Navigation	11
Understanding Site Structure Through Natural Language Processing	11
Flat vs. Deep Site Architecture	13
What to Look for After a Site Architecture Audit	15
Site Speed	19
Choosing a Better Hosting	21
Caching Method	22
Image Optimization	23
Decrease HTTP Requests	23
URL Parameters	25
Rel="canonical" Tags	27
Robots.txt Disallow	28
Redirects	28
Log File Analyzer	31
Robots.txt	32
Character Count Matters	33
Blocking SEO Bots	34
Identifying Pages Blocked by Robots.txt	34
.htaccess	35
JavaScript Crawling	36
Implementation and Testing	39
Thank You For Reading	41

Technical SEO can seem daunting to many SEOs just by the name - A LOT of SEOs would define themselves as creatives and anything to do with coding raises eyebrows.

But if you're doing SEO, at some stage or another you're going to need to work with what's going on behind the scenes.

Therefore, whether you like it or not, **you eventually have to learn about the technical aspects of your site** to help it rank even higher on Google.

Despite some people's reservations, technical SEO isn't actually all that technical. You still don't really need to know actual languages like JavaScript or PHP, and HTML is so basic my 6 year old little sister has already learned it in coding camp for kids...

If you're a complete beginner to SEO, then I recommend you check out my Learn SEO eBook first, as this one is made for those who already have an intermediate understanding of the industry and how to rank, specifically in Google.

This guide will help you understand the ins and outs of technical SEO, even if you have no idea how it works in the first place. After going through this eBook, you should have the confidence to make necessary tweaks and edits on your site, without having to ask or hire a developer every time you need to do so.

Technical SEO In A Nutshell

Technical SEO refers to the process of making your site easier to crawl and index by search engines. Nothing more, nothing less.

The entire point of having a strong technical SEO ability is to setup new sites for success and be able to work on large scale sites with hundreds of thousands (or even millions) of pages.

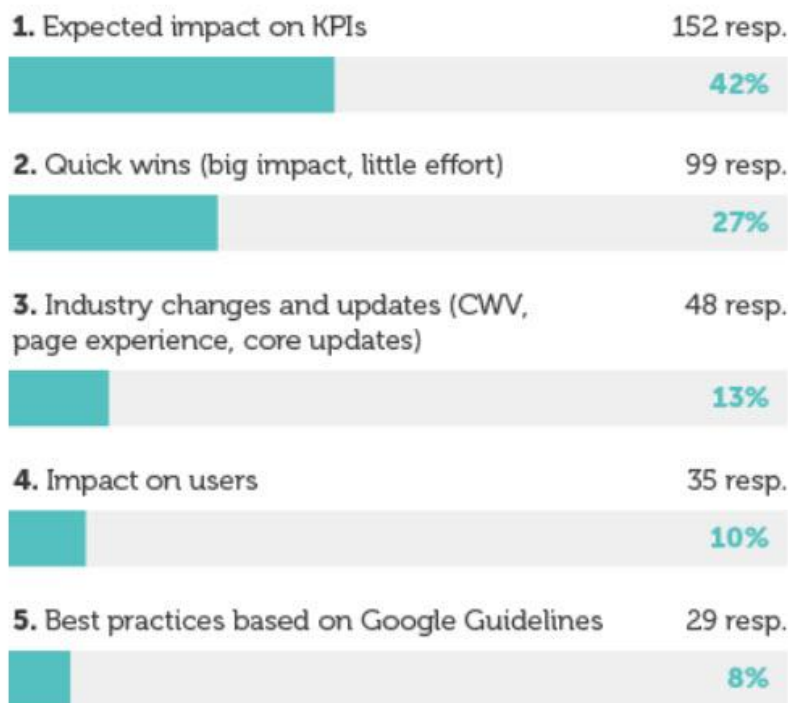
What's The Difference Between Technical SEO And OnPage?

This aspect of SEO has an overlap with on-page SEO because both mainly deal with SEO variables you can control, i.e., your website.

However, the difference between the two is that on-page SEO also considers what your target audience wants to see from your site. In contrast, technical SEO is purely concerned with getting the robots and spiders that crawl your website and report back to Google.

Now, some of the things you'll do under technical SEO could positively impact user experience. But that's more of the byproduct of optimizing the technical aspect of your website and not the end result.

A study conducted by Aira and Women in Tech SEO shows that 42% of respondents believe in the same thing -



Why Is Technical SEO Important?

The importance of technical SEO can't be overstated. Think of it as the activity that separates the good sites from the *great* ones. The former may have great content and authoritative backlinks, but it's all for nothing if search engines can't properly read these sites.

As mentioned, technical SEO is best performed even before proceeding with your content and link building campaigns. You want your site to be built on solid foundations. Analyzing your site for technical SEO opportunities gives you insights on how to maximize the results of your SEO efforts.

This means you can get the most out of the content you'll publish and the links you'll build moving forward.

So How Does Technical SEO Work?

To fully understand and appreciate technical SEO, you need to know the different parts that assemble it.

We already talked about the concept of crawling and indexing your site pages on search results. While the process seems simple enough, a lot goes under the hood that site owners and even most SEOs miss out on.

So let's break down the terms mentioned above first into their very core principles, so you have a better idea of how your technical SEO impacts the way search engines look at your site.

Search Spiders

The ones doing the crawling are actually search spiders. Also known as web crawlers, these spiders are bots that search the entire web for different URLs to store and index in their respective search engine database. URLs come in the form of site pages, images, and uploaded files and documents.

Crawling

The concept of crawling in SEO terminology comes from the ability of spiders to walk on the threads of the web they shoot.

Each search engine has its spiders to build its database of pages to index. However, Googlebot is the only one that genuinely matters due to Google being the most used search engine on the planet.

In the context of the internet, the spiders can travel from URL to URL using the links on a page. By following the internal and external links from a crawled and indexed web page, spiders can find new pages to add to their databases.

This is why *internal links* on a page are vital in helping search spiders crawl your site pages properly. Linking to your latest pages is one of the best and fastest ways to get your new and old pages crawled and indexed properly by Google and others.

Many case studies have come out of the woodwork promoting the benefits of developing a link building strategy as part of your SEO campaign.

Aside from links, search spiders crawl your site pages via your sitemap. It is a file that links out to the different URLs of your site in an organized manner.

A sitemap tells Google what the most critical pages on your site are. Using the pages linked out from the file, spiders can crawl the pages you want them to find.

Crawl Budget

As powerful as Google is, it cannot pour all of its resources into crawling all of your pages all the time. In fact, some sites are crawled more often than others for various reasons. As a result, each site has an assigned crawl budget, which refers to the number of times a search engine can crawl your site in a period.

The crawl budget for each site depends on the crawl limit and crawl demand. The former tells you how much crawling a website can handle. The latter determines the value of your site based on the number of times it's updated in a specific timeframe and how popular it is.

From these factors, you can say that the crawl budget is higher for sites that draw the most traffic, are constantly updated (whether with new content or refreshing old posts), and are hosted on platforms that can handle a lot of crawl activity.

Increasing your crawl budget naturally happens over time. You can't expect your brand new website to have a considerable crawl budget, even if it's hosted on a reliable hosting provider. The fact that it's not drawing enough traffic yet means that a search engine can't justify giving your site a sizable crawl budget.

At the same time, it's easy to waste your crawl budget with a disorganized and unoptimized site. Issues like indexed thin content, broken links, pages with long loading times, and others that you can detect in your SEO audit can deplete your crawl budget.

As a result, spiders won't be able to crawl your other essential pages on time. You'll have to wait until you replenish your crawl budget before spiders get back to work on your site.

From here, you need to fix any of the on-page SEO problems your site may have to prevent wasting your crawl budget.

For instance, if you want to disallow search spiders from crawling specific pages and directories of your site, you can create a robots.txt. This file suggests spiders what they can or can't crawl on your website.

By default, a site has no robots.txt file, which prompts search spiders to crawl the entirety of your site. But with robots.txt, you can save bots the trouble of crawling the URLs and directories you specified in it. As a result, you can redistribute the crawl budget to the more important pages of your website.

Log File Analysis

To help you gain a better understanding of how often and which spiders are crawling your site, you must refer to your site's log file.

It records requests that your web server receives regarding your site. Aside from interactions with search spiders, a log file also contains data about the:

- HTTP status code of the page being requested
- IP address where the request is coming from
- Time when the request was made
- Method of the request

Aside from knowing the crawl frequency, a log file shows you which pages are being crawled unnecessarily or too frequently. This way, you know where your crawl budget is wasted so you can make the necessary changes on your site.

The real challenge with analyzing the log file is by getting the log file itself.

The method of getting access to your log file depends on how your server is set up. There are three familiar web servers where you can get the log file from: Apache, NGINX, and IIS. If you're running your website through a CDN like Cloudflare, Kinsta, or others, you will have to get it from there instead.

From here, you need to follow the instructions on how to retrieve the log files for specific web servers. Below are links to the exact process for each:

- [Accessing Apache access log files \(Linux\)](#)
- [Access NGINX access log files \(Linux\)](#)
- [Access IIS access log files \(Windows\)](#)

If you don't understand anything from the links above, it's best to get professional help and not run the risk of messing things up for your site.

Once you have access to your log file, you can run them using the different tools discussed below. From here, you can identify where the crawl budget goes and whether or not you're spending it on your site correctly.

Without an analyzer, this is what you'll see from your log file:

Refresh Copy to Clipboard

```
185.191.171.41      com - [02/Nov/2021:12:07:15 +0000] "GET
http://www.semrush.com/bot.html" | TLSv1.3 | 0.373 0.386 0.388 MISS 0 NC:000000 UP:-      HTTP/2.0" 200 15896 "-" "Mozilla/5.0 (compatible; SemrushBot/7~bl; +h

54.36.148.154      .com - [02/Nov/2021:12:07:04 +0000] "GET /
p://ahrefs.com/robot/)" | TLSv1.3 | 0.648 0.661 0.664 MISS 0 NC:000000 UP:-      ' HTTP/2.0" 200 13921 "-" "Mozilla/5.0 (compatible; AhrefsBot/7.0; +htt

35.213.183.31      .com - [02/Nov/2021:12:07:04 +0000] "POST /wp-cron.php?doing_wp_cron=1635854824.5029881000518798828125 HTTP/2.0" 200 20 "https://
/wp-cron.php?doing_wp_cron=1635854824.5029881000518798828125" "WordPress/5.8.1; https://christopherjanb.com" | TLSv1.3 | 0.250 0.262 0.261 - 0 NC:0
40000 UP:-

165.225.58.232     com - [02/Nov/2021:12:06:09 +0000] "GET /wp-content/uploads/2018/05/cropped-K_i
02-crop-32x32.png HTTP/2.0" 200 2
48 "https://i
/" "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like G
ecko) Chrome/95.0.4638.54 Safari/537.36 Edg/95.0.1020.40" | TLSv1.3 | -- 0.000 - 0 NC:000000 UP:-

165.225.58.232     com - [02/Nov/2021:12:06:08 +0000] "GET /wp-content/uploads/
HTTP/2.0" 200 4808 "http
s://
/" "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) C
hrome/95.0.4638.54 Safari/537.36 Edg/95.0.1020.40" | TLSv1.3 | -- 0.000 - 0 NC:000000 UP:-

165.225.58.232     com - [02/Nov/2021:12:06:07 +0000] "GET /wp-content/plugins/affiliatebooster-blocks/dist/assets/fontawesome/webfonts/fa-solid-900.woff2
HTTP/2.0" 200 62472 "https://
css-0b91756a46234c4a0e9f96cb3f496c6
6.css" "Mozilla/5.0 (Windows NT 10.0; Win64; x64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/95.0.4638.54 Safari/537.36 Edg/95.0.1020.40" | TLSv1.3 | -- 0.227 - 0 NC:0
```

These lines won't make sense from an SEO perspective unless you run it using a log file analyzer.

Load Balancing

Going back to web hosting, it is vital to host your website on reliable servers. There are lots to choose from out there, and you probably won't go wrong with some of the more popular choices out there, like WPX Hosting and Kinsta, to name a few. These hosting providers offer fast loading speed, excellent customer service, and outstanding performance.

However, an aspect of hosting that isn't talked about much often in the realm of SEO is load balancing.

This system refers to the ability of the server infrastructure to spread out the requests across multiple servers. It prevents specific servers from getting overloaded with requests, which is common among high-traffic websites.

More importantly, load balancing improves the overall performance of your website on the webserver. It enables you to utilize the server resources you have at your disposal.

In the context of SEO, load balancing may cause problems to your user experience. Because it constantly utilizes multiple servers during a user session, your site cannot build session persistence. This term refers to the act of directing requests from the same client, whether it's a visitor or search spider, to the same backend server. Now, if the client session takes more than a few minutes, expect load balancing to have distributed the requests to different servers, which slows down the user experience.

An example of this is when users are filling out a multi-step form on your site. Instead of loading all the steps in a single server to improve loading speed and performance, the web host may delegate the user requests to different servers. While the host may think it's doing you a favor, it's slowing down the request time to process the information.

At worst, visitors will get so annoyed by the lag in loading your site pages that they'll be forced to leave your site.

In this case, you can refer to your log file to review the request your load balancer received for your server. This way, you can analyze how it managed the requests and uncover insights to help you approach load balancing.

Indexing

Finally, after discussing the different factors that could affect your site's crawlability, we move on to indexing its pages.

Once the page is saved in its database, the bots index the pages for their most relevant keywords.

In this part, your on-page SEO game comes into play. How optimized and useful your content is for the keywords it's targeting will determine how high or low it can rank on SERPs. The more

valuable the page is concerning your target keyword, the higher the chances it can rank on search results.

Of course, there's more to ranking web pages than just on-page SEO. But technical SEO has made getting your site pages indexed and ranking on Google possible. It's just a matter of developing SEO campaigns to grow its authority and achieve high rankings moving forward.

Tools We'll Need

There's a lot to take in from the technical SEO process discussed above. Part of this is implementing the operation of getting your web pages crawled and indexed properly.

Since technical SEO requires multiple steps to implement correctly, you'll need the help of tools to collect and gather data for analysis. It would be next to impossible to analyze your website on a technical level without using software to automate some processes.

However, when choosing which technical SEO tools to use, make sure that it meets specific criteria to ensure that you gather the correct information about your site.

Below are some of the considerations when choosing a technical SEO tool to use:

- **Site-wide auditing** - The tool should enable you to analyze the site based on various factors affecting technical SEO performance.
- **Page-wide auditing** - The platforms must have the ability to inspect pages and get granular with each one. This helps you identify key elements that keep the page you're analyzing from ranking on the first page or help maintain its position at the top.
- **JavaScript analyzer** - Google can render and execute JavaScript (JS) as part of the crawling and indexing process. Using a tool's JS crawler, you can see how it affects page speed, code coverage, and accessibility.
- **Logfile analyzer** - To help you make sense of the log file, you need to run it using a log analyzer that most technical SEO tools have. It should reveal to you crawl budget waste and identify fake bots that simply scrape your site and hurt your server's performance, among other things.
- **Integrations** - You want a tool that will help you incorporate data from third-party sources like Google Search Console and Google Analytics. Even better, these tools should help make SEO data easy to understand so clients and stakeholders can better handle how the site's performance is impacting their bottom line.

With these considerations in mind, below are tools that possess the majority of the features above and that you should choose as your technical SEO platform moving forward:

- [Screaming Frog SEO Spider](#) - Arguably the most popular SEO audit tool, Screaming Frog is a desktop-based platform capable of auditing websites on a page level and providing SEO recommendations for each based on the gathered data.
- [Sitebulb](#) - Another desktop-based application, Sitebulb is a more streamlined version of Screaming Frog, making it much easier to use for beginners. It offers hints to users on how to approach specific site or page issues. There are also tutorials helping you how to execute the recommendations so you can fix the problems yourself.
- [DeepCrawl](#) - This tech SEO tool is geared towards enterprise-level websites with thousands—if not millions—of pages that require auditing. Aside from its advanced auditing features and customized reporting, Deepcrawl allows you to compare audits over time using the historical data view to help you determine issues that still need fixing.
- [JetOctopus](#) - JetOctopus is a sleek cloud-based auditing tool that provides insightful data organized according to priority, so you know which issues to deal with first. You can also stream live logs so you can see client requests as it happens in real-time. It also integrates seamlessly with your GSC and GA.
- [SEOTesting.com](#) - Aside from being a website audit tool, SEOTesting.com is the only one on the list with extensive testing features so you can maximize your SEO efforts. Choose from split testing, time-based tests, or URL-group-based testing so you can put your campaigns and ideas to the test.

All of the tools above require a monthly or annual subscription, although most have free trial periods to help you test-drive the software and decide if it's the one for you. Screaming Frog SEO Spider is the only one with a free tier that only allows you to crawl up to 500 URLs.

If you're looking for a free solution, [Ahrefs Site Audit](#) is a fine choice. While it may be as feature-rich as the ones above, you can use it on your websites for free, provided that you can prove ownership of these sites. Plus, it offers more than decent information about websites once audited.

Factors to Analyze in Technical SEO

When starting with any of the technical SEO tools above, you will have to run an audit of your site to unpack any of the issues it may have. However—and this is where things get tricky—a part of the audit will cover on-page SEO problems and recommendations.

Earlier, I mentioned the overlap between tech SEO and on-page SEO, which will be evident once you see the audit reports. While treating your on-page problems is just as important as fixing your site's technical issues, this guide is focused on the latter.

So, if you want more information on optimizing your website using data from an audit, I highly recommend that you get a copy of my [Omniscient OnPage](#) guide. It covers a gamut of topics

that aren't covered here, such as content and topical optimization, E-A-T, CRO for SEO, and others.

Below are things you must look out for after auditing a site from your technical SEO process. And for checking these factors, we'll be looking at different tools mentioned to help you unearth these issues. I'll show you a few tools you can use to help speed up the process about implementing the fixes.

Site Structure and Navigation

We talked about how search spiders can crawl your site pages through internal links on your content as well as your sitemap. However, it's not as simple as randomly linking out to different pages on your site just for the sake of making the process of crawling them much more accessible.

For instance, linking pages dealing with different topics together may make crawling these pages much more effortless. However, they won't benefit your SEO in the long run because linking the pages adds no value to the functionality of your site.

In other words, **your internal links must make sense to spiders**. This is where information architecture (IA) or site architecture comes into play.

IA is not visible to users, unlike the site navigation (UI). Instead, only spiders can view them based on organizing your pages according to topic or taxonomy.

Understanding Site Structure Through Natural Language Processing

The goal of IA is to create topic relevance to your niche. Linking related pages together helps spiders pick up the relationship among interlinked pages based on the words or "entities" used in these pages.

The ability of spiders to acknowledge entities in pages comes from the natural language processing (NLP) model it uses to organize and understand the context and intent of their content.

If you go to [Google Natural Language AI](#), you can take its Natural Language API demo for a spin. Paste a paragraph on the text box before clicking on "Analyze."

DEMO

Natural Language API demo

Try the API

Try the API

Google, headquartered in Mountain View (1600 Amphitheatre Pkwy, Mountain View, CA 940430), unveiled the new Android phone for \$799 at the Consumer Electronic Show. Sundar Pichai said in his keynote that users love their new Android phones.

ANALYZE

[See supported languages](#)

Once analyzed, you will see a list of entities organized according to schema and salience.

Entities	Sentiment	Syntax	Categories
<Google>₁, headquartered in <Mountain View>₂ (<1600 Amphitheatre Pkwy, Mountain View, CA>₁₂ <1600>₁₆ <Amphitheatre Pkwy>₇, <Mountain View>₂, <CA 940430>₈ <940430>₁₄), unveiled the new <Android>₃ <phone>₅ for <\$799>₁₃ <799>₁₅ at the <Consumer Electronic Show>₁₁ . <Sundar Pichai>₄ said in his <keynote>₉ that <users>₆ love their new <Android>₃ <phones>₁₀ .			
1. Google Wikipedia Article Salience: 0.19	ORGANIZATION	2. Mountain View Wikipedia Article Salience: 0.18	LOCATION
3. Android Wikipedia Article Salience: 0.14	CONSUMER GOOD	4. Sundar Pichai Wikipedia Article Salience: 0.11	PERSON
5. phone Salience: 0.10	CONSUMER GOOD	6. users Salience: 0.09	PERSON
7. Amphitheatre Pkwy Salience: 0.07	LOCATION	8. CA 940430 Salience: 0.05	OTHER

Instead of simply relying on links to determine the search ranking of the pages it indexes, Google now banks on NLP to help gain a deeper understanding of the content on each page. This way, it can identify terms closely related to your target keywords and make the connection from there.

If the linked pages mention similar or related entities, the correlation among these pages becomes much higher. If you do these on all pages on your site, you can strengthen your site's topic relevance to its niche, resulting in higher positions on SERPs for your keyword.

Flat vs. Deep Site Architecture

For publication sites, IA can be simplified into topic clusters. This process involves grouping together pages exclusively talking about their respective topics relevant to your niche. For example, you can create a WordPress category about a topic in your niche and publish all articles related to that topic. Do the same thing to the other categories you have on your site relevant to your industry.

On the other hand, e-commerce sites naturally have a more complex IA assuming that a site sells different types of products. Due to the breadth of niches in an online store, you must organize the pages into multiple topics and subtopics, making it easier for spiders to crawl and index them.

The difference between publication and e-commerce sites is the former has a flat site architecture while the latter has a deep site architecture.

Flat Site Architecture



Deep Site Architecture

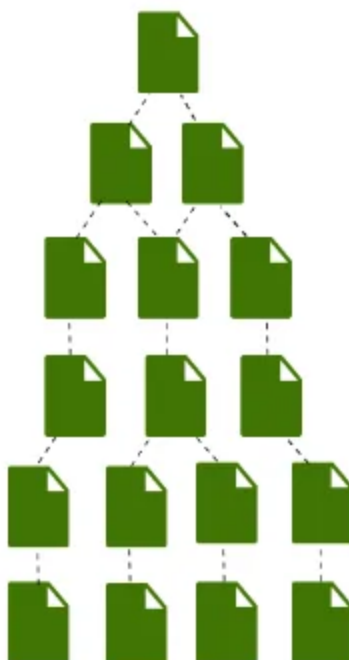


Image source: [SEOQuake](#)

Ideally, a 2-3 level site architecture is ideal so spiders won't have trouble crawling and indexing the site. This means all pages on a site are two to three clicks away from the homepage.

E-commerce sites will be hard-pressed to achieve a site architecture of this level due to the subtopics they will have to create to organize their pages and content appropriately.

As mentioned, you can't view IA unless you have an audit tool that will help you visualize the structure. Screaming Frog has this feature—click on Visualizations > Crawl Tree Graph to open a new window and see how spiders view your website.



I've blurred the URLs of the pages, but it should be clear to you that this blog has only two levels in its site architecture. It is shallow enough for spiders to index the pages without any problems.

Again, things get more complicated when dealing with e-commerce sites due to most of them having subtopics within subtopics. This could push their site architecture to at least four levels deep, which is not suitable for SEO.

What to Look for After a Site Architecture Audit

Once you have information regarding your site architecture from your chosen audit tool, you need to find a way to trim down the levels to the lowest possible and make crawling your website much more manageable.

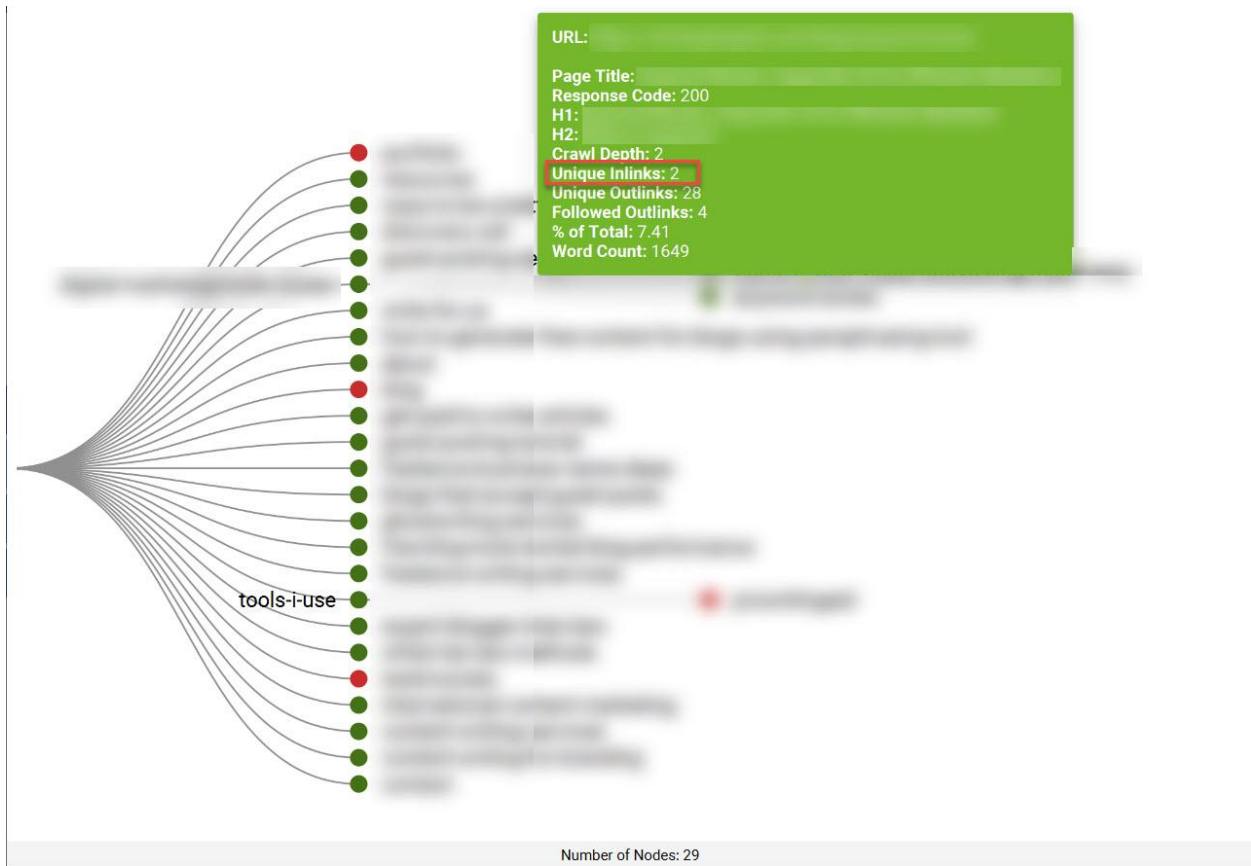
You can do this by bringing the subcategories closer to your homepage using your site navigation.

The solution on how to help improve your site architecture depends on the problems your site has. But below are some common fixes you can implement, especially for enterprise-level or e-commerce websites:

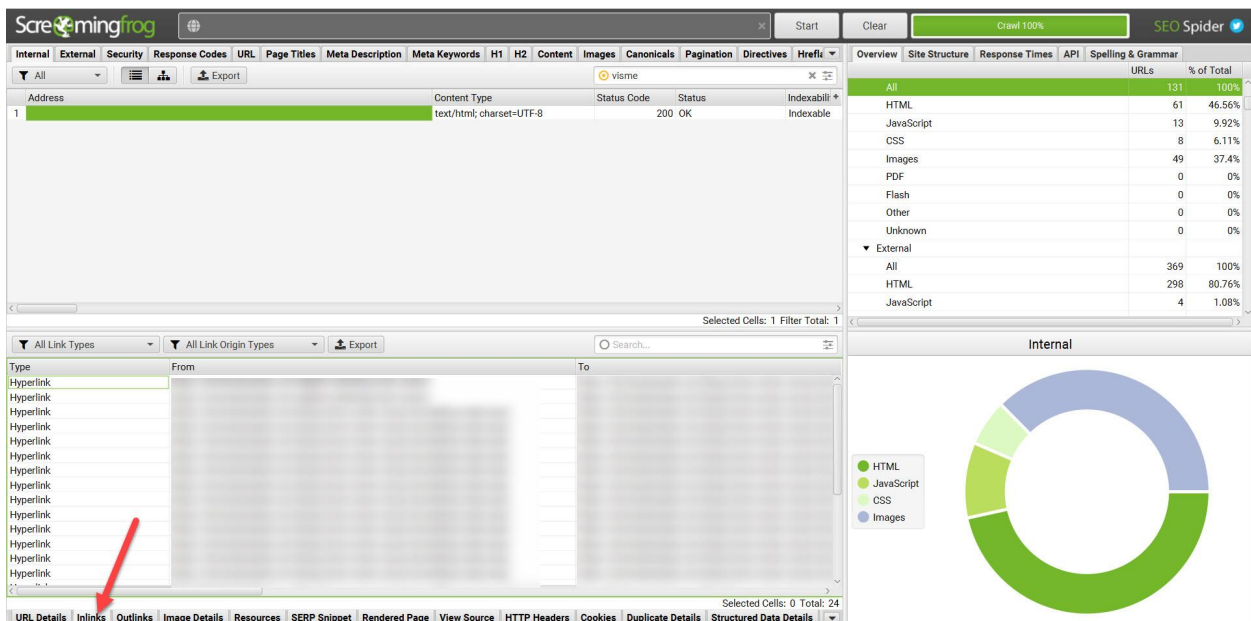
- **Cross-link among silos** - Ideally, you want to contain topic clusters within themselves to help improve your site's topic relevance. Then again, this is SEO we're talking about—there's no one-fit-fits-all answer to specific problems. At the same time, there is no rule regarding internal links when building content silos. Google would most likely adhere to linking to internal pages should make sense to users and search spiders. In cases when trimming your site to three levels or less is not possible, you can link pages from one silo to a page from another silo when possible, i.e., if the pages talk about the same topic despite being in different silos. Try to find pages away from the homepage that you can link to from pages near the homepage. Doing this could improve the crawlability of pages that are four levels or higher away from the homepage.
- **Leverage your blog for internal linking opportunities** - If you have existing blog posts, find pages away from the homepage that you can naturally link to in your posts. If not, develop a blogging strategy to create content that will link to these pages. Assuming that these are product and category pages, you need to create content with commercial keyword intent that allows you to organically mention your products and categories. At the same time, consider adding blog posts optimized for informational intent keywords to help balance the commercial content in your site.
- **Organize site navigation** - Links within the content are what Google value in terms of crawling and indexing. If you can't link pages to the content, there are other places where you can link out, such as your top navigation bar, sidebar, and footer section. Then link to the critical pages on the top navigation, the next important pages on the sidebar, and the least important ones on the footer. Keep in mind that links on the footer are mostly devalued. But for user experience, consider utilizing the footer to showcase links you want search spiders to crawl if you can't link to these pages on the content body. Finally, if you have multiple links to the same page, bots will consider the first links they see on the page according to the source code. This is referred to as the First Link Priority. So, if you link to the same product page on a page thinking that it'll improve crawlability, it won't. Spiders will only consider the anchor text of the link that it reads first on the page. Keep this in mind when building links in your navigation sections.

If you run your site using Screaming Frog and check its Crawl Tree Graph, things will be easier to analyze. But implementing the fixes, you'll find here is done manually.

If you want to point more internal links to a page that's far away from the homepage, you can highlight it on the Crawl Tree Graph view and see its information.

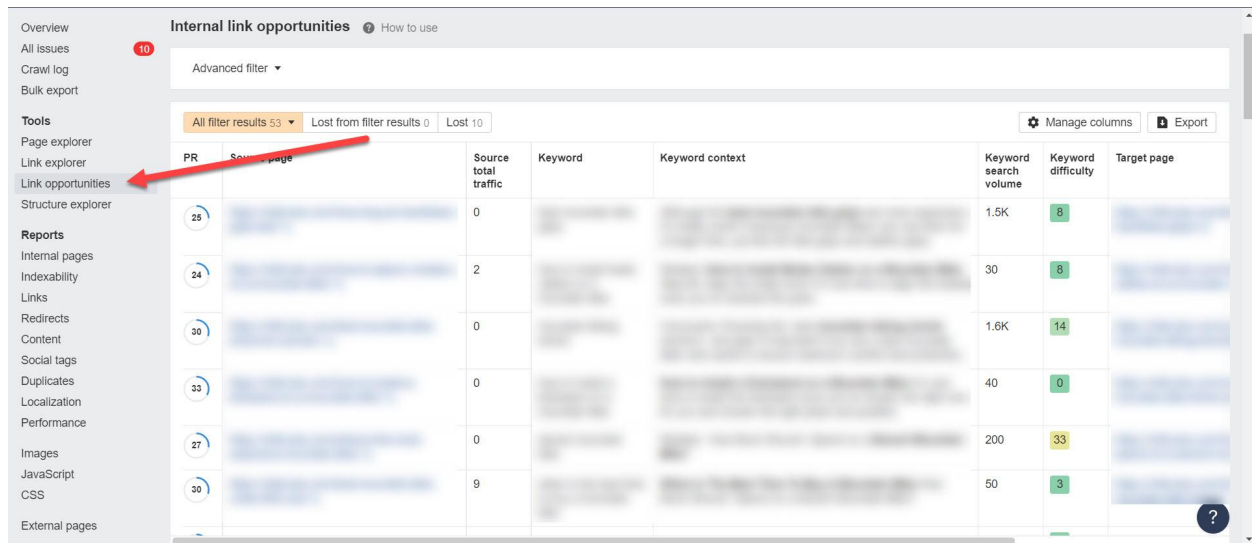


You can identify the links pointing to the page by looking for the URL of the page you want to analyze and clicking on the Inlinks tab at the bottom.



From here, you can decide if the pages linking to the target page are good enough to keep. If you don't have enough internal links to the target page, you need to search for pages to link to it. This is the real challenge because you must manually find the most relevant pages that are the most pertinent to the target page. At the same time, you have to figure out how you plan on including the link in the page and what anchor text to use.

Thankfully, some tools could help speed up the process. One of them is Ahrefs' free Site Audit. Upon verifying that you own the site, you wish to analyze, run a site audit and wait for it to complete. Once done, go to the results and click on Tools > Link opportunities to view internal link opportunities on your site.

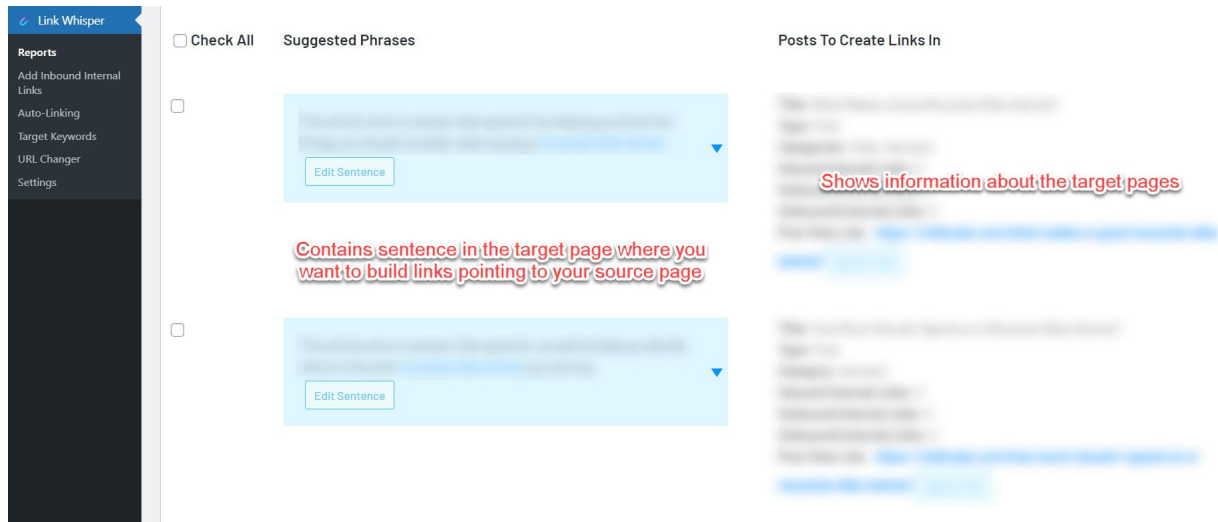


PR	Source page	Source total traffic	Keyword	Keyword context	Keyword search volume	Keyword difficulty	Target page
25		0			1.5K	8	
24		2			30	8	
30		0			1.6K	14	
33		0			40	0	
27		0			200	33	
30		9			50	3	

The page shows you the different source pages you can link to and the anchor texts you can use that are already mentioned on the target pages. While the results you'll see here aren't perfect, they nonetheless provide you with ideas as to which pages you should consider linking out to.

If you're using WordPress, a premium plugin you can use to help you automate internal link building is [Link Whisper](#). Upon installing and running a scan of your site, go to the page that you want to build more links to and run a report for it.

On the next page, you will see a list of target pages where you can link your source pages to multiple target pages with just a click of a button.



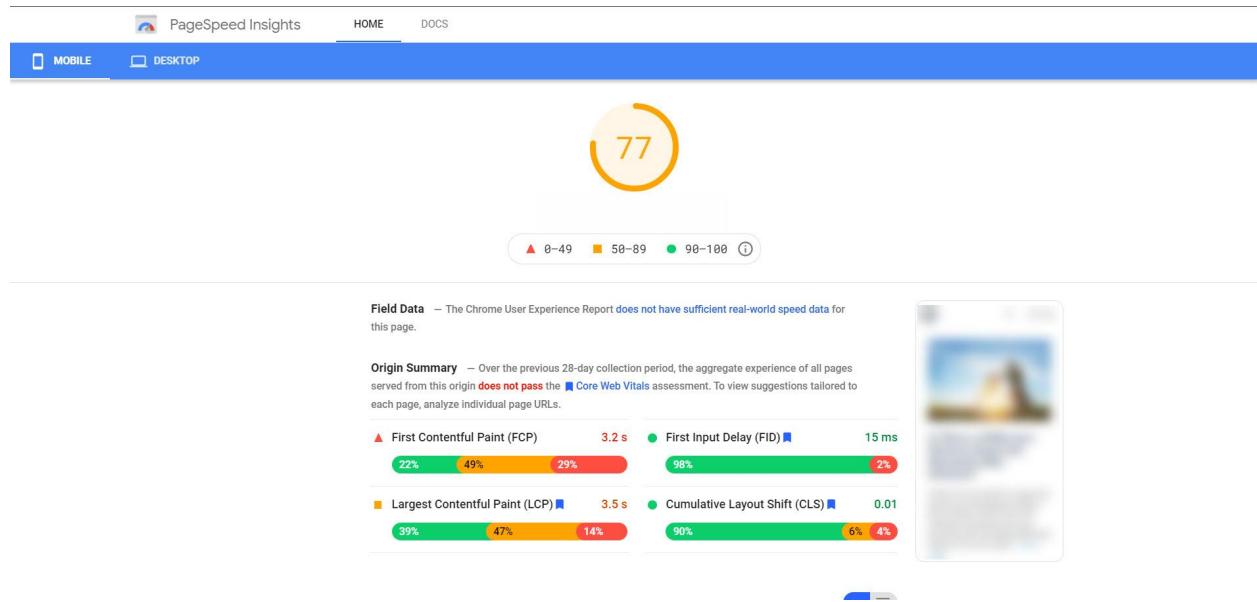
Each suggestion uses an anchor text that already exists in the target pages to your source page. If you're not satisfied with the suggested sentence and anchor text, you can edit it from the page. Then tick the boxes where you want to include your target page, and the plugin will publish the links for you.

Again, while Link Whisper won't always provide you with the anchor text you want for the source page, you can at least edit them on the target pages. This makes the job much easier for you to help increase crawlability by linking to your source page from other target pages.

Site Speed

A lot has been said about site speed. Ever since Google finally acknowledged [site speed as a ranking factor in 2018](#) and with the [Page Experience Update in 2021](#), more and more people are trying to decrease loading speed by increasing the efficiency of loading resources on your site.

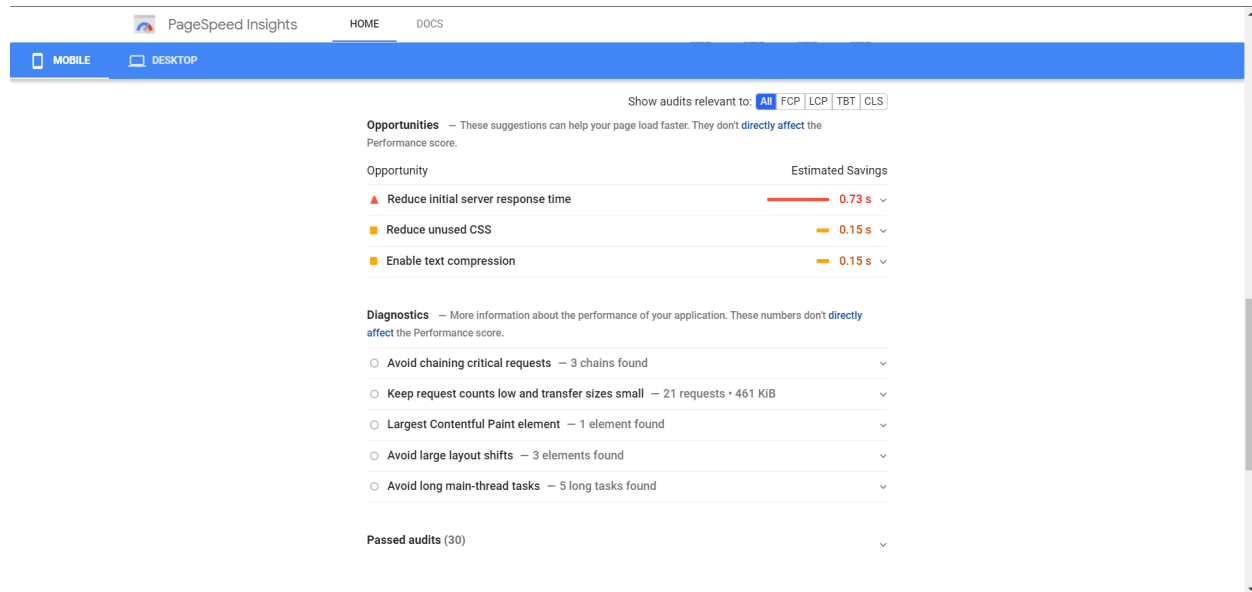
You can view how your site is performing by running a test on [Google PageSpeed Insights](#).



The results of your URL analysis will show you the Core Web Vitals (CWV) scores of your site. The tool breaks down your score into the following metrics:

- **Largest Contentful Paint (LCP)** - tells you which elements on the page's main content have loaded.
- **Cumulative Layout Shift (CLS)** - lets you know how much of the content has shifted its position up or down due to certain elements loading on the screen.
- **First Input Delay (FID)** - measures the delays in executing a particular action performed by users, such as clicking on a button.

Scrolling down the page, you will see opportunities and diagnostics detailing how to improve the site's performance.



While the data you'll find here is helpful in understanding how to implement the changes, they won't mean a thing to a person who's not a web developer. Some of the issues and solutions you'll see in this part requires coding knowledge and experience to eliminate the problems.

However, just because you don't know much about web development doesn't mean you won't be able to do something to improve your site's CWV scores. The bulk of the work you can do to hike up your site's performance in terms of loading speed and efficiency boils down to the following factors:

- Hosting
- Caching method
- Image optimization
- Decrease HTTP requests

Thanks to plugins, applying the changes to these factors would be much easier if you have a WordPress site. And since most website owners are using WordPress as their CMS, we'll detail how you can make the changes yourself below.

It's possible to enforce the same principles discussed below on non-WordPress sites, but you may need to rely on alternatives o successfully implement them.

Choosing a Better Hosting

We've touched upon the topic of web hosting earlier regarding log file analysis and load balancing. In this case, you need a web hosting that will not only do the following but also observe the things below for your site:

- **Speed** - How fast or slow your site loads depends on whether you're using shared or dedicated hosting and the server specs used to run your website.
- **Uptime** - A hosting using advanced technology as its servers ensures that your site is up and live for everyone to access for as long as possible with minimal downtime.
- **Server location** - You want a web host with servers located nearby your target audience and primary visitors. This helps load your site pages faster due to the proximity of the server to your visitors.
- **Customer support** - If you encounter issues with your website, you want someone from the hosting company to attend to your concerns ASAP. The longer your customer support is unresponsive to your problems or provides insufficient answers to your queries, the longer your site will be down for your visitors to see. That means decreased trust from your audience and lost revenue.

If you search for the best web hosting platforms for your website, you'll be greeted with different suggestions from different people.

Part of the reason is that some hosting companies offer high commission rates for every successful sale an affiliate makes. That means people are referring you to certain hosting platforms is because they want to earn money and don't care if the hosting is good.

This is common among affiliates, so it's best to be wary of who to trust when asking for suggestions for reliable and fast web hosting.

For instance, if you trust [Matthew Woodward](#) and [Matt Diggity](#), both will tell you that [WPX Hosting](#) is the best one out there. According to them, it offers the best and fastest web performance compared to others.

You can do your research and dig deeper if you want to use a different hosting provider. But keep the factors in mind if you want to find the right hosting for you.

Caching Method

In a nutshell, web caching allows your site to generate static HTML versions of your pages and save them on your servers. Whenever a visitor requests the page, instead of reaching out to the hosting to load its resource-intensive version, it will be the generally lighter HTML version for faster and more efficient loading.

This process may sound complicated, but you can get a caching plugin for all your pages automatically. Aside from caching, you also get the following features that help make loading your site much easier for servers:

- **Lazy loading** - You can delay loading your images and videos until visitors scroll down to the part where the media appears.

- **Database cleanup** - It can scan your database for unnecessary files that take up space and can be deleted.
- **CDN integration** - Connect your website with a content delivery network (CDN), so you can increase your site's loading speed by fetching data from nearby data servers.

Among the different caching plugins available for WordPress, [WP Rocket](#) is one of the best and most popular of the bunch. It possesses all the core features above and is easy to set up and use.

Free alternatives to WP Rocket include [WP Super Cache](#) and [W3 Total Cache](#). Both are capable of caching your site pages effectively, but both lack the features and simplicity that WP Rocket has.

Image Optimization

If you have an image-heavy site like an online store, you need to ensure that product photos are optimized and compressed to their smallest file size without compromising quality.

Usually, you would have to compress your images before uploading them to your site. However, the problem lies with having hundreds of unoptimized photos on your site. Do you have to download the images, optimize them, and reupload them to help improve your site's loading speed?

Thankfully, you don't have to do that if you're running a WordPress site. With premium image optimization plugins available, you can command it to decrease the file size and retain the image quality of existing images on your database. At the same time, all new images you upload are automatically optimized, so you don't have to.

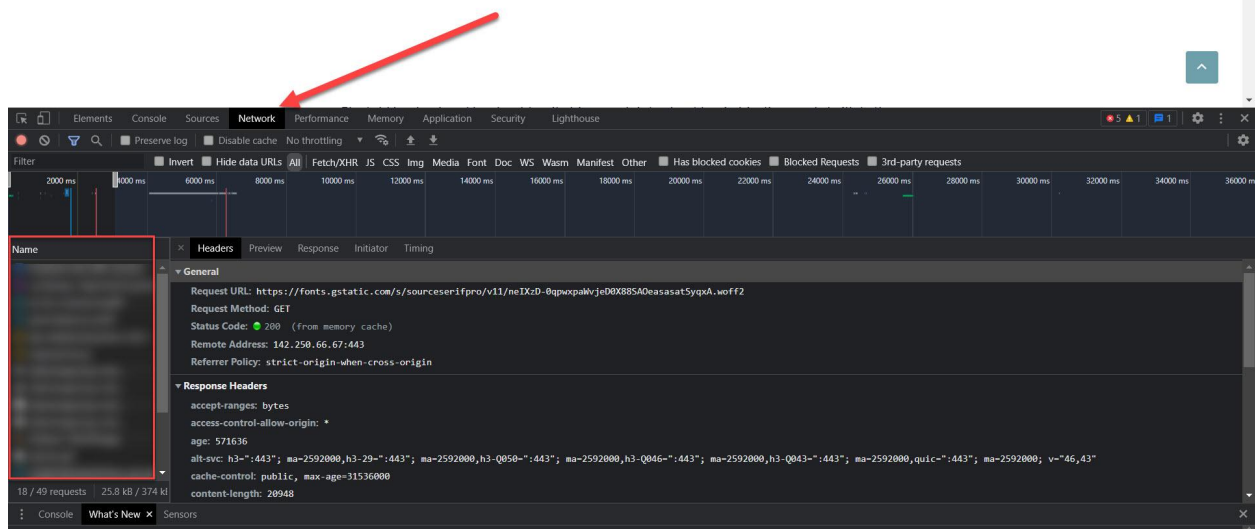
One of the most popular ones in the market is [ShortPixel](#), although you really can't go wrong with other choices such as [EWWW Image Optimizer](#) and [Smush](#).

The great thing about ShortPixel and Smush is the WebP support. You can convert all your images into the WebP format, which offers superior image compression among image formats.

Decrease HTTP Requests

If you downloaded many plugins for your site, remember that some of them are operating in the background, even if you're not using them for a particular page.

To view the scripts of plugins running on the page on Google Chrome, right-click on the page and select "Inspect" to open the developer tools. From here, click on the Network tab and reload the page to view the scripts and files here.



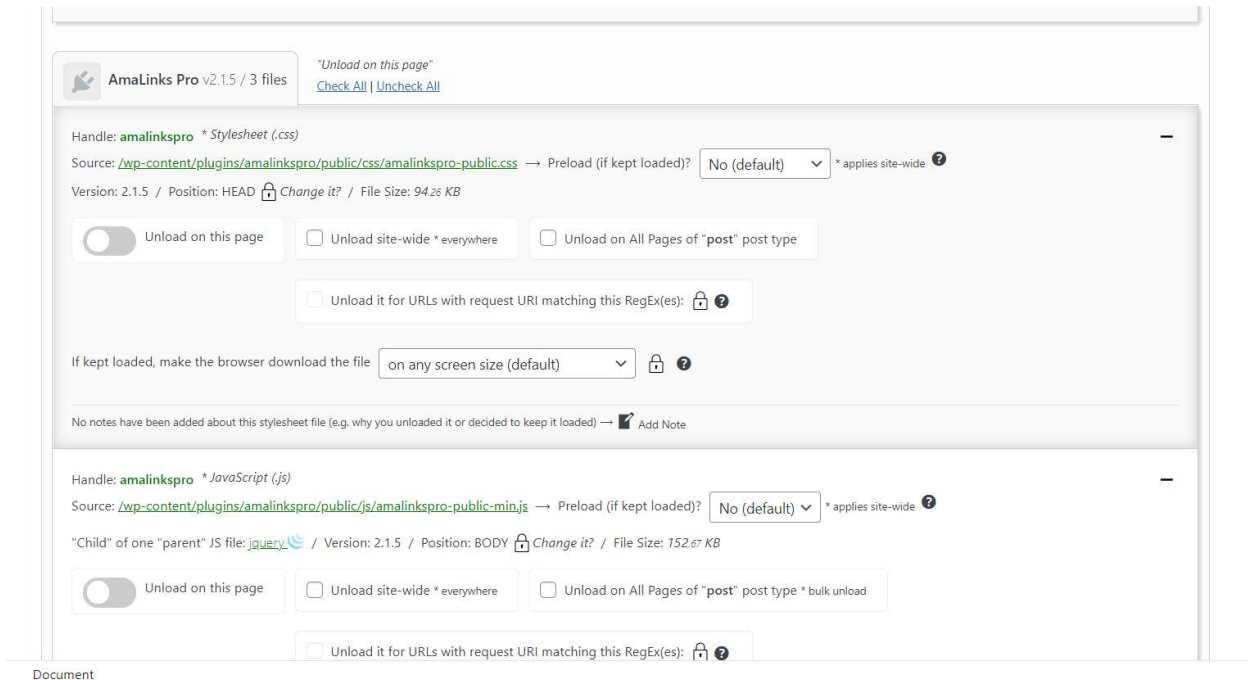
Now, it will be difficult to find out which scripts or plugins you're not using for this page but are running here.

Using a caching plugin to minify and concatenate JS and CSS files should help some of the problems you'll encounter here. Another solution is to delete plugins that you don't use or replace them with leaner and more lightweight plugins.

However, if you remove or replace these plugins but are still left with a sizeable amount of plugins that run in the background, you need to turn them off manually.

You can do this by downloading and installing the [Asset Cleanup](#) plugin. It has a free version which is good enough to use alongside WP Rocket. Another plugin you ought to consider using is [Perfmatters](#). It's essentially the same as Asset Cleanup but with a different and arguably better interface. Its only downside is that Perfmatters isn't free. In this case, Asset Cleanup should be more than enough to do the trick.

Once activated, go to a page on your site that you want to analyze. In the editor, scroll down the page until you see this section:



From here, you can unload plugins that you're not using on the page. This way, users can load the pages efficiently without having to worry about unnecessary scripts.

When unloading plugins and scripts, make sure the page is not using them. If you accidentally unloaded a script, it might cause problems on the page. It's best to review the page after unloading a script to verify that no problems are found afterwards.

URL Parameters

URL parameters are strings of characters added to a URL for tracking and analytics purposes. Using a [URL Campaign Builder](#), you can identify where the traffic came from in a particular campaign.

The screenshot shows the 'Campaign URL Builder' tool in Google Analytics. At the top, there's a header with the Google Analytics logo and 'Demos & Tools'. Below this is the title 'Campaign URL Builder'. There are three tabs: 'WEB' (selected), 'PLAY', and 'IOS'. A descriptive text states: 'This tool allows you to easily add campaign parameters to URLs so you can measure Custom Campaigns in Google Analytics.' Below this is a section titled 'Enter the website URL and campaign information' with a note: 'Fill out all fields marked with an asterisk (*), and the campaign URL will be generated for you.' The form contains four input fields: 'website URL *' (with a hint: 'The full website URL (e.g. https://www.example.com)'), 'campaign ID' (with a hint: 'The ads campaign id.'), 'campaign source *' (with a hint: 'The referrer (e.g. google, newsletter)'), and 'campaign medium *' (with a hint: 'Marketing medium (e.g. cpc, banner, email)').

URL parameters are also generated when users are filtering results of your online store using:

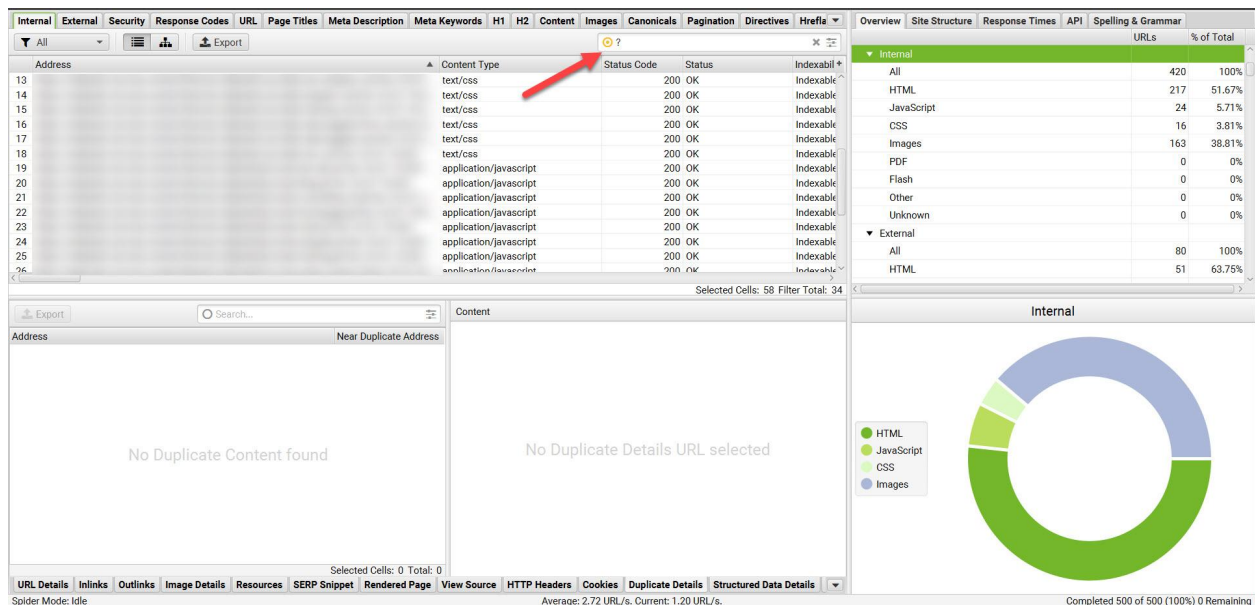
- faceted navigation
- multiple pages displaying your posts on a blog category
- and search queries made using your site's native search feature.

As helpful as URL parameters are for these reasons, they are an SEO nightmare.

The biggest reason why URL parameters suck from an SEO standpoint is that search spiders view them as individual URLs. Even if the URLs are pointing to the same page, spiders will still view each URL as a page of its own.

As a result, you're creating [duplicate content](#) that not only wastes your site's crawl budget and creates keyword cannibalization issues.

To identify URL parameters that create SEO issues for your site, run Screaming Frog and enter "?" on the search bar once the audit has concluded. It will filter pages that contain "?" in the URL.



Ideally, you purchase a license key of Screaming Frog to be able to extract all URL parameters.

Keep in mind that searching for “?” would also return pages that aren’t necessarily URL parameters like JavaScript, CSS, and others. But you can organize the content type so you can isolate the pages with multiple URL parameters.

To also know how Google views these URLs of yours, you can use search operators such as the one below:

`Site:yoursite.com inurl:"parameter string"`

Replace yoursite.com with your actual site and “parameter string” with the string of characters that you think could be causing duplicate content. Here, you should know the exact parameter strings your site is using on these pages.

If you have lots of URL parameters for a page on your site, it’s time to put a stop of time before they make matters worse.

Rel="canonical" Tags

One of the most straightforward fixes to this problem is to identify the canonical URL of a page. If you’re using a plugin like Rank Math, you can indicate the canonical URL on the editor for that page.



This does add a `rel="canonical"` tag for that page—all URL parameters of the page will point back to it. As a result, you can consolidate all ranking factors to a single page and eliminate duplicate content.

On the downside, spiders will still crawl its URL parameters, so you're still wasting the crawl budget in this sense.

Robots.txt Disallow

To fix this issue, add this line onto your robots.txt:

```
Disallow: /*?*
```

What happens here is that search spiders won't crawl URLs on your site with a "?" in them.

This solution complements the `rel="canonical"` tag fix perfectly. While this solution won't help consolidate ranking signals from the different URL parameters into a single page, the `rel="canonical"` tag fix will.

Redirects

Redirection indicates a change of a page to another location.

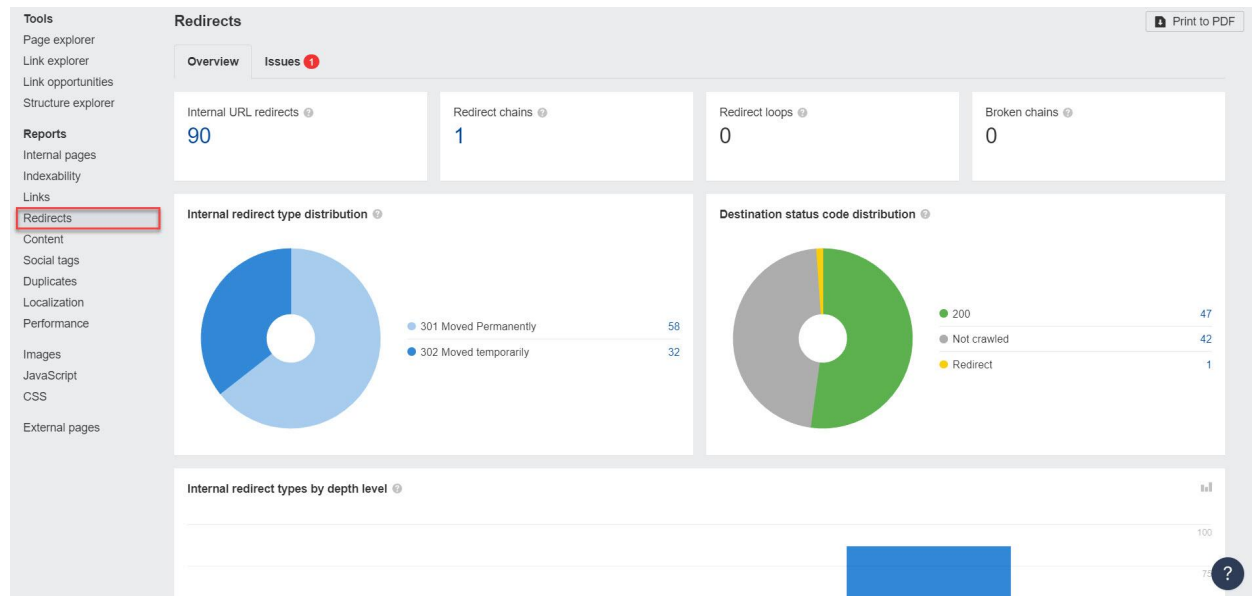
Two of the most common response codes for redirection are 301 and 302. The former informs spiders and bots that the change is permanent. All of the page's ranking signals also go to its new location.

When using 301 for redirection, make sure that you've decided to do so. If you change your mind and revert to the previous location, the page may no longer rank anymore for its keywords. In other words, there's no way to undo the effects of 301 redirections.

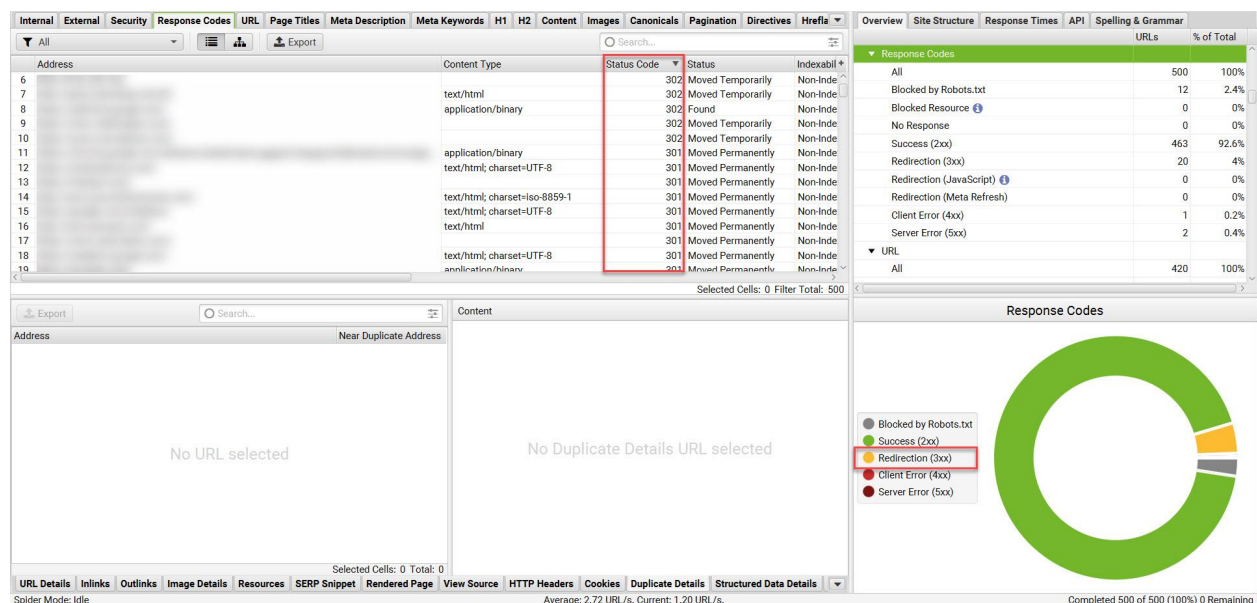
302, on the other hand, is a temporary redirect. This means that while clients are brought to the page's new location, they are informed to refer to the original location for future requests.

302 is identical to 307—the difference is that the 302 code suggests that they go to the new location. If they decide to stay in the old location, it's not considered an error. This reason is why some would advocate for the use of 307 instead of 302 status code.

To identify which pages on your site have these response codes, you can view them on Ahrefs Site Audit by going to Reports > Redirects.



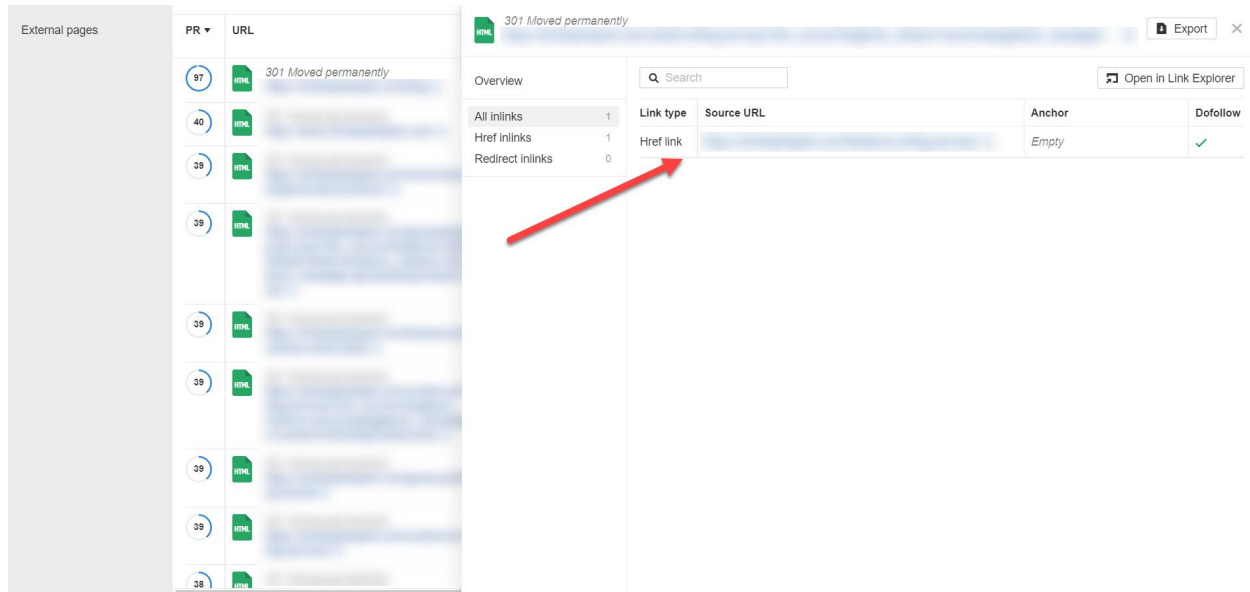
You can also view these status codes on Screaming Frog after the audit process.



From here, the issue isn't necessarily with the redirects. Instead, it has to do with the pages linking to old locations of the redirected pages.

By linking to the old pages, you dilute the link juice to flow among the pages of your site because of the additional requests caused by the redirects. From a user perspective, redirects slow down the loading speed of the pages, which could lead them to exit your page out of frustration.

Using Ahrefs, identify pages linking on the original locations of the redirected page.

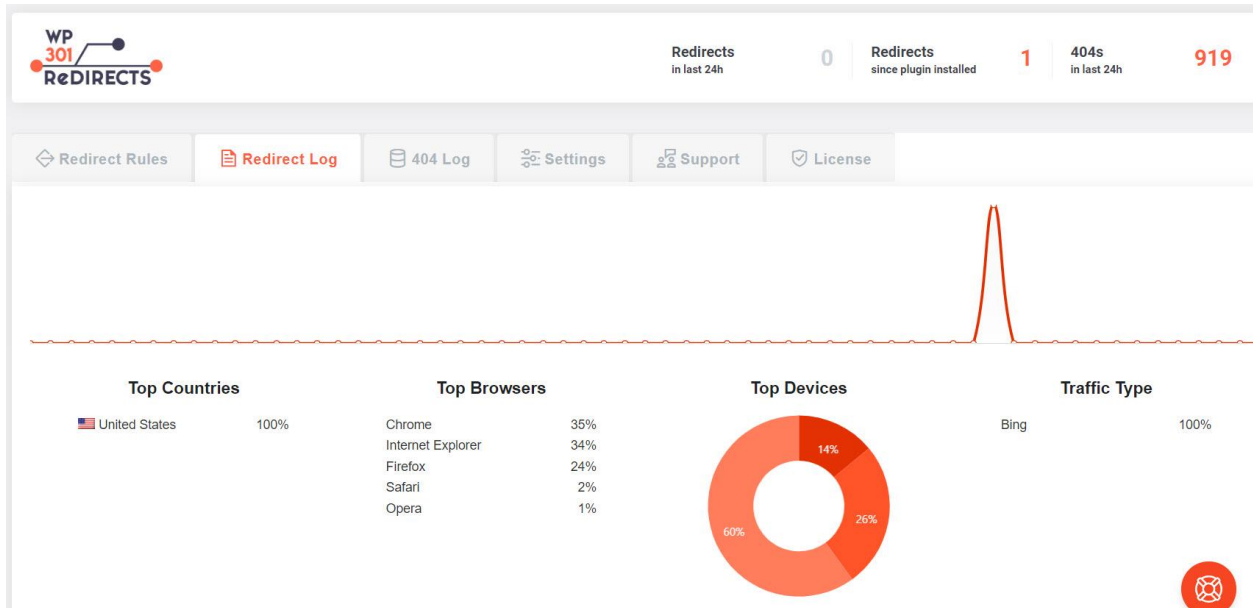


The screenshot shows the Ahrefs Link Explorer interface for a page with a 301 Moved permanently status. On the left, a list of external pages is shown with their PR (Page Rank) and URL. The main panel displays the 'Overview' section, which includes a search bar and a table of incoming links. A red arrow points to the 'Link type' column in the table, which shows 'Href link'.

Link type	Source URL	Anchor	Dofollow
Href link		Empty	✓

You then have to manually edit the link on your site to link to the new location instead of the old one.

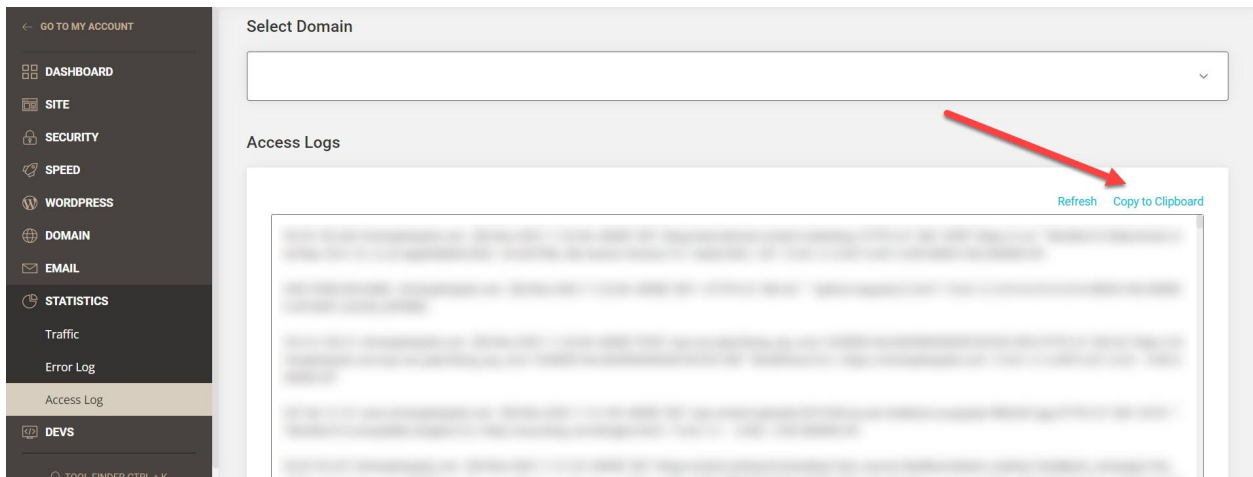
On WordPress, you can also keep a redirect log to see which old URLs of redirected pages are getting the most traffic. From here, you can determine why this is the case and implement the necessary fixes to each. A plugin like [WP 301 Redirects](#) ought to do the trick.



Log File Analyzer

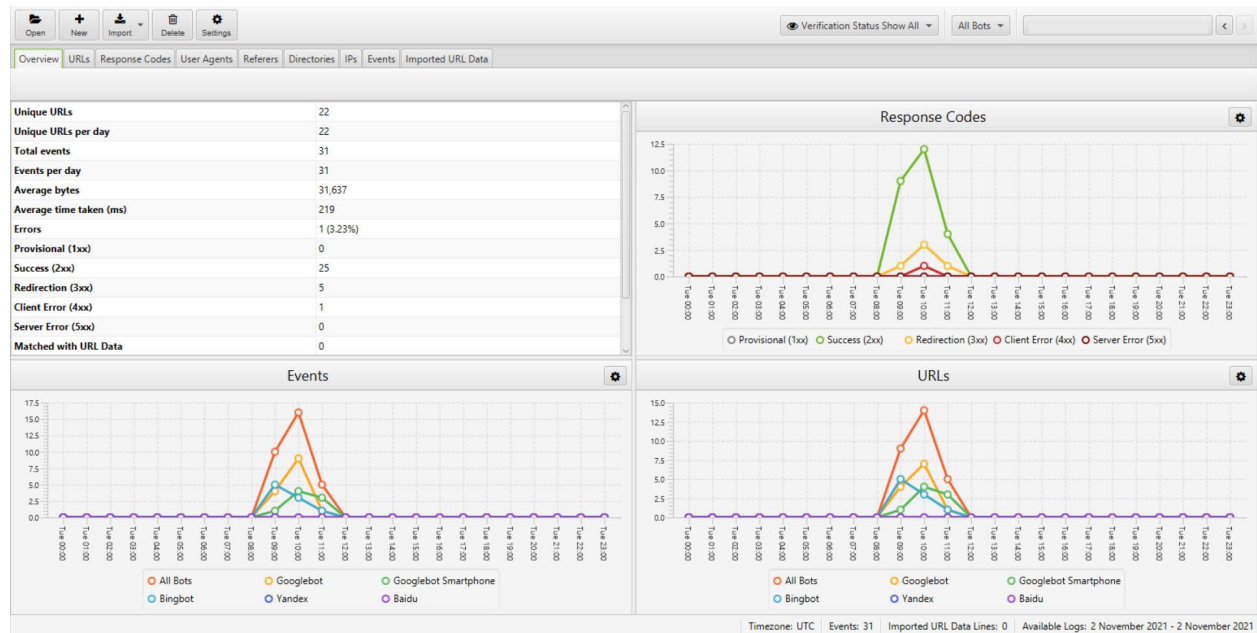
We've already discussed analyzing your log file to review your site's requests from clients. To see these requests, you need a log file analyzer to help you make sense of the log file you will have retrieved from your server.

We linked to the different ways you can download the log file from your server. However, there are other ways you can get the log file from your hosting, especially if you want to analyze the current requests. For example, if you're using Siteground, go to Site Tools > Statistics > Access Logs, copy the logs to a clipboard, and save them in a notepad using .log as the file extension.



From here, you can upload them to your Screaming Frog Log File Analyser software. It's a separate tool than the SEO Spider—you must [download it here](#).

Upload the file and wait for a few seconds before you see something like this:



You can then identify the bots crawling your site and which pages are crawled the most. From here, determine which among the unimportant pages are hogging your site's crawl budget and which pages need some love from spiders.

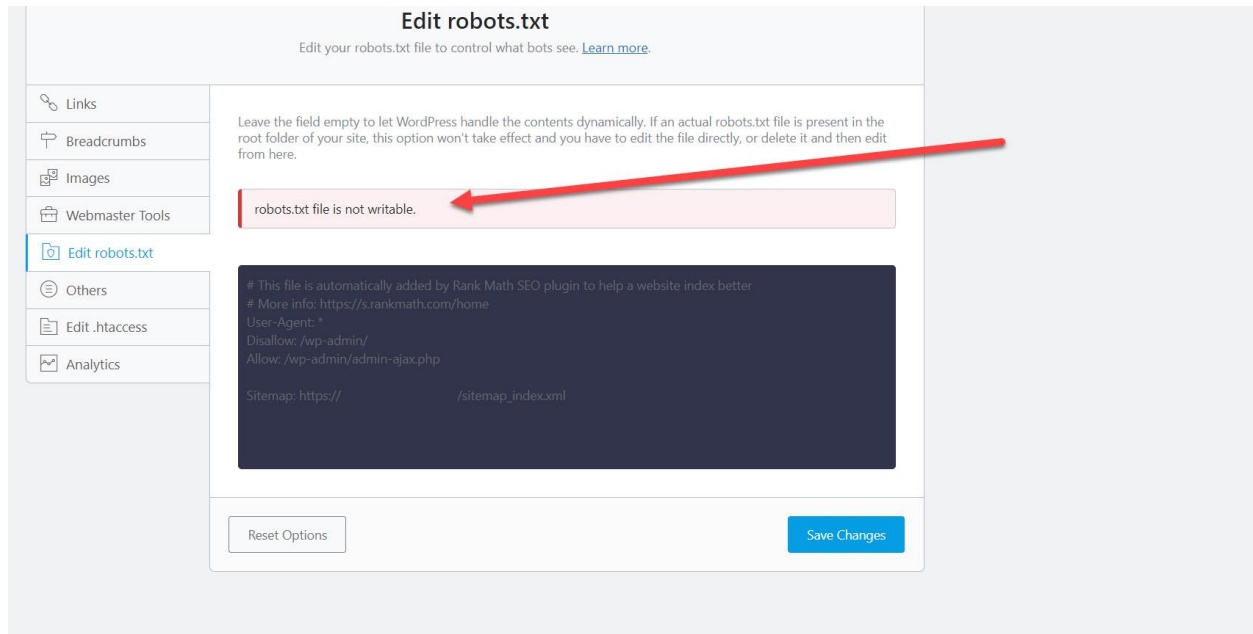
Robots.txt

We mentioned earlier that a robots.txt file could help improve your site's crawl budget by disallowing search spiders from crawling URLs of your site.

However, a potential issue that may arise from using robots.txt, for this reason, is that you may have disallowed URLs that you want spiders to crawl by accident. This could cause a significant impact on your SEO performance, which is why you want to review and analyze your robots.txt. Doing so helps guarantee that all vital pages on your site are crawled and indexed correctly.

Initially, you can edit your robots.txt from your hosting's Control Panel or FTP. You can then allow and disallow pages to be crawled by editing the file from there.

If you have Rank Math, you can edit robots.txt from your WordPress dashboard. This makes making changes much easier since you can manage the file directly from your site.



If you have an existing robots.txt before setting up Rank Math on your site, you won't be able to edit it from WordPress. To fix this, delete the current robots.txt on your host and let Rank Math create one for you.

When editing your robots.txt, below are the best practices you should observe. They may not apply to your site, but it's best to keep them in mind moving forward:

Character Count Matters

Google and Bing search spiders consider the length of the URL or folder to be allowed or disallowed into consideration. Here's an example of a robots.txt:

```
User-agent: *
```

```
Allow: /category/subcategory/
```

```
Disallow: /category/
```

Here, search engines are unable to access the `/category/` directory except for `/category/subcategory/`.

However, the rules change when the robots.txt is written like so:

```
User-agent: *
```

```
Disallow: /category/
```

```
Allow: /category/subcategory/
```

Using the directives here, Google and Bing are the only search engines allowed to access the `/about/` directory, which includes `/about/company/`. The reason is that the `allow` directive has more characters than the `disallow` directive.

Keep this rule in mind when editing your `robots.txt`. How you organize the directives can alter how bots crawl your website.

Blocking SEO Bots

Some site owners block bots of popular SEO tools like Ahrefs and SEMrush using `robots.txt`. The reason is to prevent them from wasting your crawl budget.

This is a problem most encountered by large or popular websites in which their competitors are analyzing their sites using these tools. You can check your log file to see how many requests you're getting from these bots. If the requests are higher than usual, consider blocking the bots of tools you don't use for developing and building SEO strategies and campaigns.

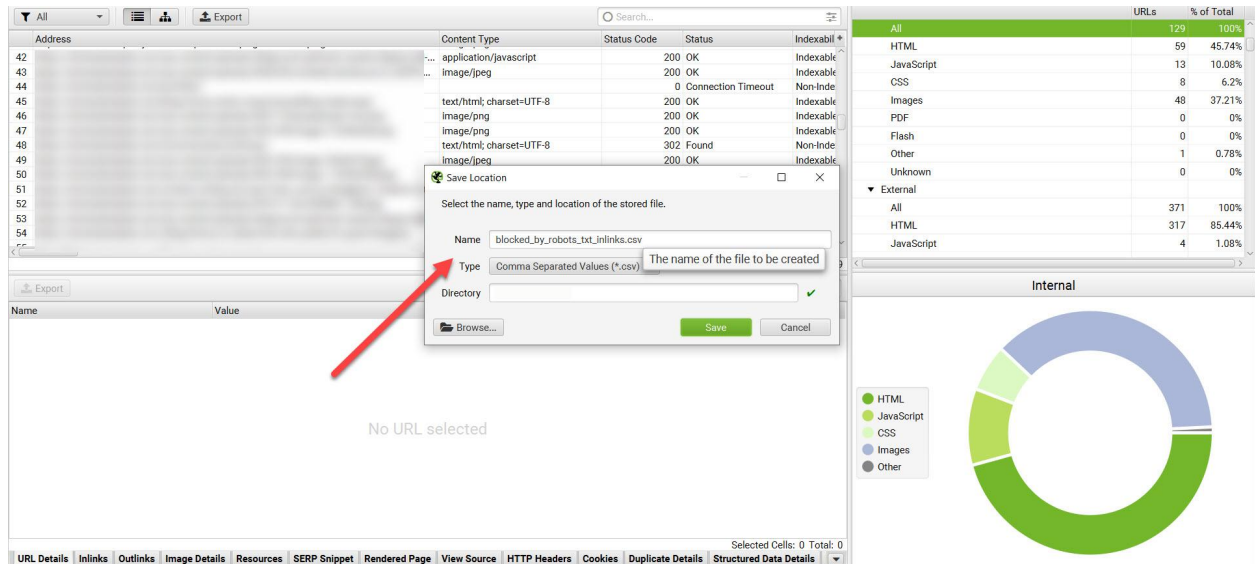
Also, private blog networks (PBNs) disallow these bots from crawling their sites to prevent people from detecting patterns in their link-building initiatives.

But for small to medium website owners, there's no reason to block them for now unless they're disrupting your site's performance.

Identifying Pages Blocked by Robots.txt

Aside from editing your `robots.txt`, you must first identify pages that are blocked by the file from crawling. This enables you to determine URLs that shouldn't be blocked so spiders can crawl them properly.

Using Screaming Frog, you can check the URLs with "Blocked by robots.txt" under response codes. You can even export the list of files here for your reference or if you want to include it in a separate report.



Once you have edited and fixed issues on your robots.txt to resolve these issues, you can run another scan on Screaming Frog to verify the changes.

If you have a license of Screaming Frog, you can also edit robots.txt from the software, so you don't have to go back and forth from editing and checking the crawled pages.

.htaccess

.htaccess is a file that allows you to configure the files in your main servers in how search spiders and bots will treat them. Implementing the best practices of .htaccess optimization makes crawling your site easier for spiders, resulting in better SEO performance.

Below are things that .htaccess can do for your site if configured correctly:

- **Create SEO-friendly URLs** - If you have essential pages using the URL parameters and want to edit them all at once instead of fixing each page, you can write a code on .htaccess to help you rewrite these URLs.
- **Take complete control of editing URLs** - Some CMSs don't allow users to change the page's URL before publishing. If that's the case, you can turn to .htaccess to override the CMS and make site-wide implementations.
- **Handling 40x pages** - Instead of allowing visitors to access an error page on your site, you can redirect them to its working version by configuring the file correctly.
- **Hotlink prevention** - If you have images on your site, some site owners copy them without saving them on their servers. As a result, they're using *your* servers whenever the images load on *their* site. To prevent this from happening, configure your .htaccess to restrict them from doing this.

You can do other things with .htaccess, such as caching for site speed, managing redirections, and more. You can do these things using the different plugins mentioned above for WordPress.

In most cases, you're better off not touching .htaccess unless there's no other choice than to do so. But if you're not using WordPress to run your website, tweaking your .htaccess is your alternative.

However, make sure to save a copy of the original .htaccess file before you start editing it. If you input incorrect code, you can easily break your site without knowing why. If this happens, just recover the old version of the file to restore the website.

JavaScript Crawling

Dealing with JavaScript (JS) on your site pages can be daunting since most scripts are heavy and resource-intensive.

Back then, Google was unable to crawl websites built on JS until Google Caffeine in 2010. However, instead of crawling them immediately, Google bots crawl the scripts on the second wave of indexing. In the first wave, the spiders only crawl HTML and CSS since they are more lightweight and thus easier to load.

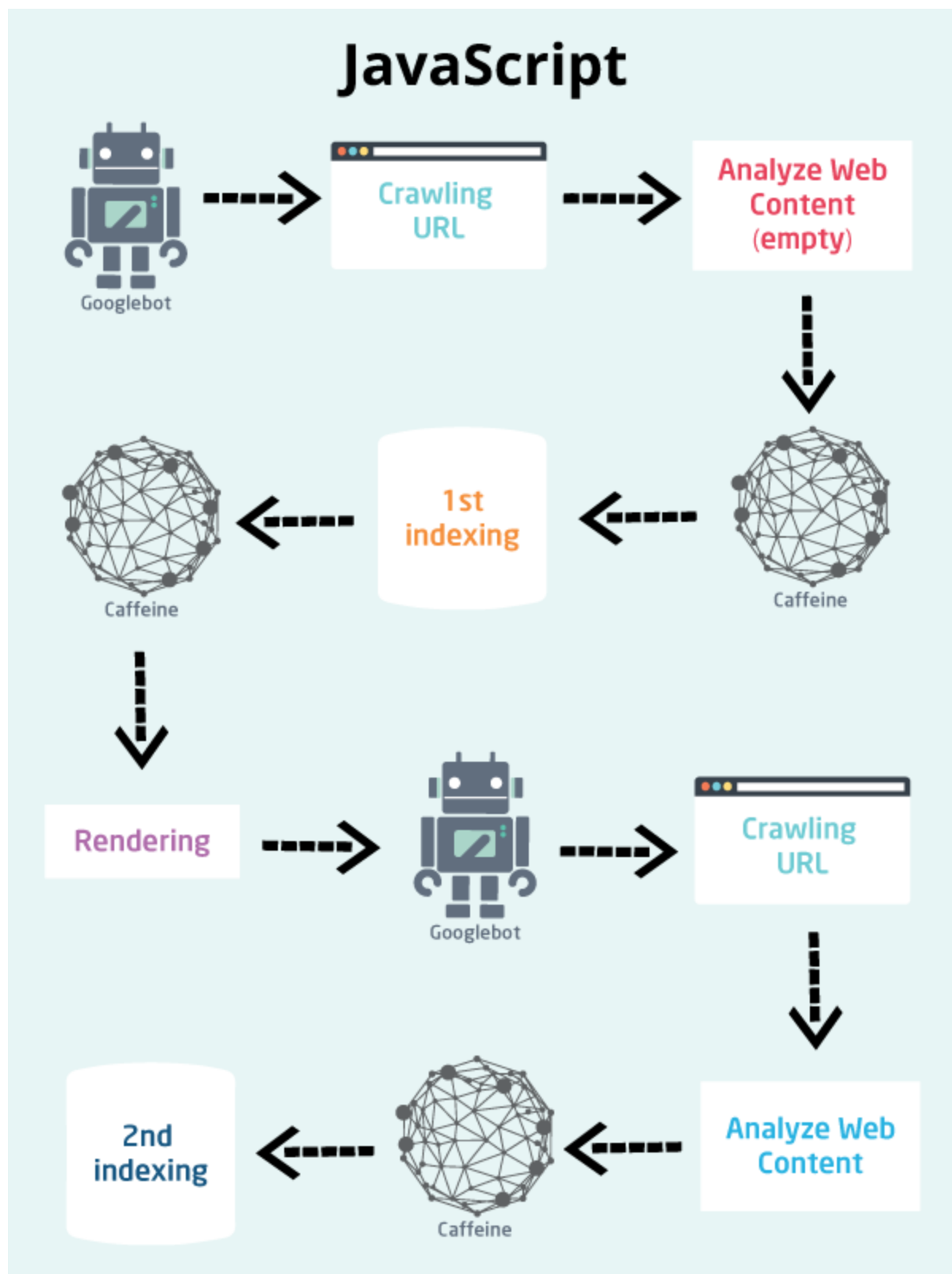


Image source: [SEOPressor](https://seopressor.com)

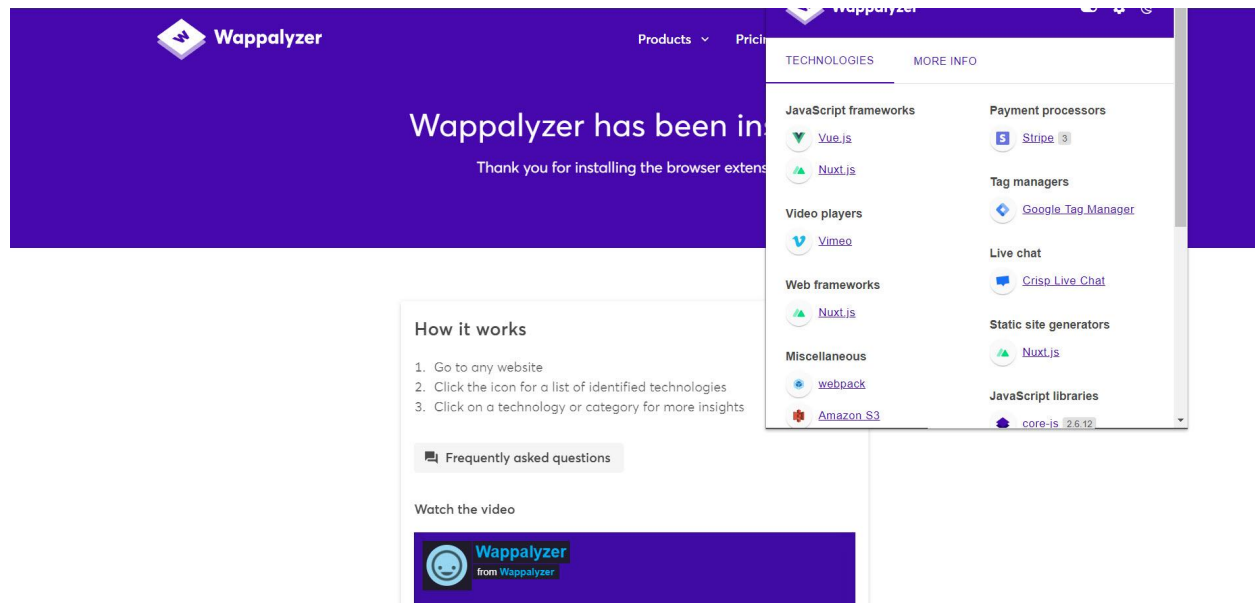
Knowing how and why JavaScripts are crawled and indexed this way allows you to understand the significance of using them on your website. While ranking JS-heavy site pages is possible, it depends on how you implement these scripts as part of the page.

For instance, the page's JS version may not match its cached version on search engines. Due to the discrepancy, Google bots won't process the page, leading it to not ranking on search results at all.

Also, not all search spiders can read scripts properly. Therefore, expect JS-heavy websites do not rank well on other search engines compared to simple, static HTML sites.

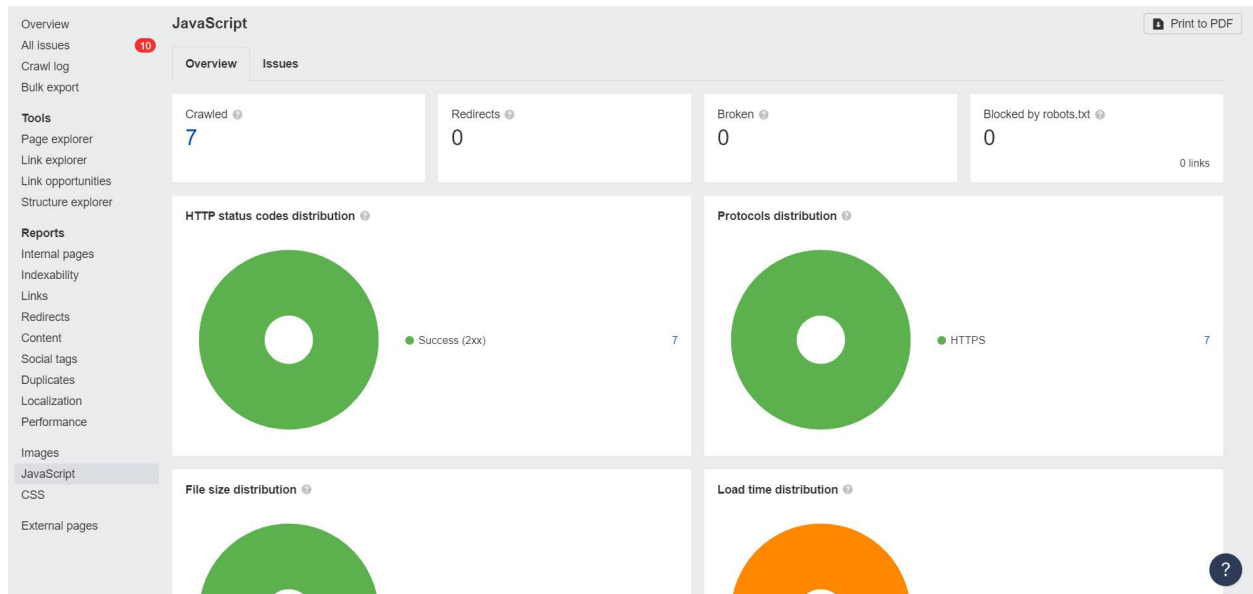
Again, the use of JS is acceptable if there is no choice but to keep them on your page. However, as much as possible, limit the use of scripts sitewide so you don't experience the problems that these scripts may produce on your site.

When auditing your site for JavaScript, identify first the scripts you're using for your site. Download the [Wappalyzer](#) to determine the scripts your website is using to run it.



From here, you should have the big picture of what to watch out for when conducting a JS site audit.

To ensure that all JavaScripts are working correctly on your site, you can check their response rates on Ahrefs after a site audit.



One way to determine some of the issues you will find in your audit is by using another Chrome extension called [View Rendered Source](#).

[View rendered source](#) URL: <https://www.jonhogg.com/view-rendered-source/install.html> Tip: Rendered view may also show page changes made by other Extensions

Show/hide: ☒ Raw ☒ Rendered ☒ Difference Bugs, comments? DMs open: [@Itshogg](#)

Raw	Rendered	Difference (0 of 376)
<pre><!DOCTYPE html> <html> <head> <meta name="robots" content="noindex"> <!-- Global site tag (gtag.js) - Google Analytics --> <script async src="https://www.googletagmanager.com/gtag/js?id=UA-47933396-1"></script> <script> window.dataLayer = window.dataLayer []; function gtag(){dataLayer.push(arguments);} gtag('js', new Date()); gtag('config', 'UA-47933396-1'); </script> <title>Untitled Webpage</title> <script src="somescript.js" defer=""></script> <link rel="stylesheet" href="style.css"> </head> <body> </pre>	<pre><!DOCTYPE html> <html> <head> <meta name="robots" content="noindex"> <!-- Global site tag (gtag.js) - Google Analytics --> <script async src="https://www.googletagmanager.com/gtag/js?id=UA-47933396-1"></script> <script> window.dataLayer = window.dataLayer []; function gtag(){dataLayer.push(arguments);} gtag('js', new Date()); gtag('config', 'UA-47933396-1'); </script> <title>View Rendered Source installed</title> <script src="somescript.js" defer=""></script> <link rel="stylesheet" href="style.css"> <style> @font-face { font-family: "wticons";</pre>	<pre><!DOCTYPE html> <html> <head> <meta name="robots" content="noindex"> <!-- Global site tag (gtag.js) - Google Analytics --> <script async src="https://www.googletagmanager.com/gtag/js?id=UA-47933396-1"></script> <script> window.dataLayer = window.dataLayer []; function gtag(){dataLayer.push(arguments);} gtag('js', new Date()); gtag('config', 'UA-47933396-1'); </script> <title>View Rendered Source installed</title> <script src="somescript.js" defer=""></script> <link rel="stylesheet" href="style.css"> <style> @font-face {</pre>

It compares the raw source code of the page to the rendered page by Google and shows the elements that didn't load correctly. You can then spot the complex scripts that are not loading on the rendered version.

Implementation and Testing

Once you have identified your site's technical issues, it's time to resolve them *one at a time*.

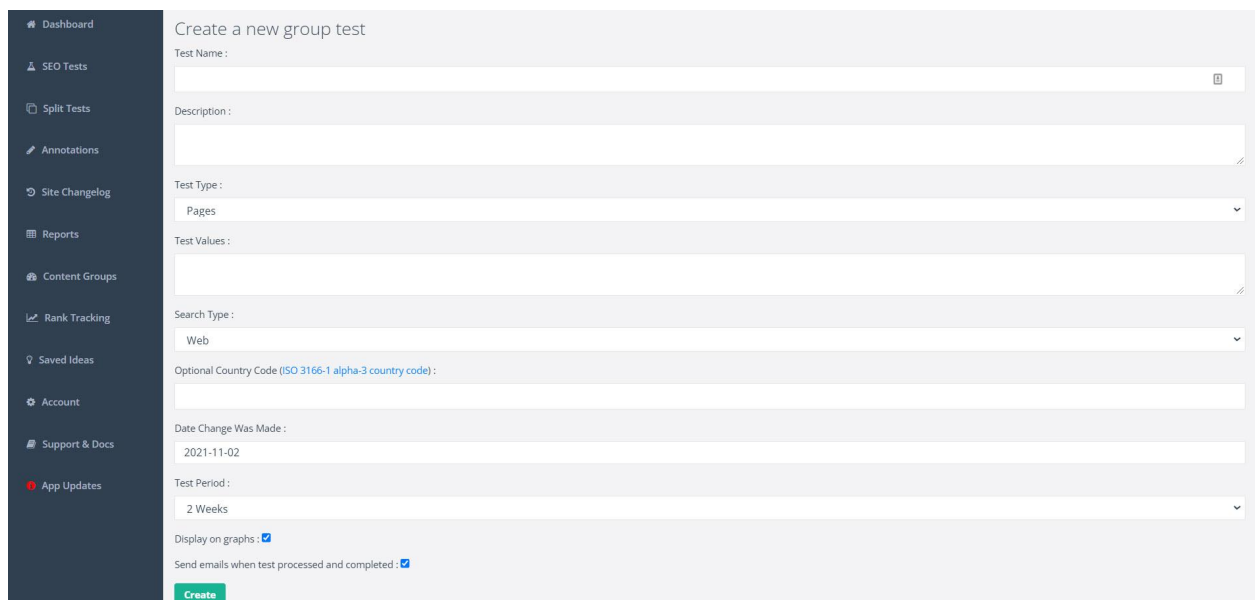
In SEO, it's not intelligent to make page and sitewide changes all at once. To be fair, it depends on the issue—if the problem is putting a damper on your rankings, it's best to attend to these issues immediately.

However, if you want to test certain ranking factors and determine which among them holds more weight than the other, then you need to try one element at a time.

By isolating the issue, you focus on the controlled variables and keep uncontrolled variables at bay. This enables you to verify if the changes you made to the variable being tested helps in improving your site rankings.

However, keep in mind that it's best to test elements from pages that aren't ranking on the first page of SERPs. You don't want to put top-ranking pages to the test only for them to drop out of the top spots because the test backfired.

Using a tool like SEOTesting.com, you can organize your tests into single tests (testing a specific URL after implementing a change), group tests (testing groups of URLs), or URL redirection tests.



The screenshot shows the 'Create a new group test' interface on SEOTesting.com. On the left is a dark sidebar with navigation links: Dashboard, SEO Tests, Split Tests, Annotations, Site Changelog, Reports, Content Groups, Rank Tracking, Saved Ideas, Account, Support & Docs, and App Updates. The main form area has the following fields and options:

- Test Name :** A text input field with a copy icon.
- Description :** A larger text input field with a copy icon.
- Test Type :** A dropdown menu currently set to 'Pages'.
- Test Values :** A text input field with a copy icon.
- Search Type :** A dropdown menu currently set to 'Web'.
- Optional Country Code (ISO 3166-1 alpha-3 country code) :** A text input field.
- Date Change Was Made :** A text input field containing '2021-11-02'.
- Test Period :** A dropdown menu currently set to '2 Weeks'.
- Display on graphs :** A checkbox that is checked.
- Send emails when test processed and completed :** A checkbox that is checked.
- Create** : A green button at the bottom left of the form.

You can describe the tests and set the duration of the date to have ample time to collect the data and allow the results to occur.

From here, you can verify if the changes you made have a positive or negative effect on its search rankings.

Thank You For Reading

I hope you enjoyed this eBook, it's created with MANY, MANY years of research and testing.

Once you have a complete understanding of technical SEO, you should be able to build pitch-perfect websites from scratch.

Don't forget you can also check out all of [my other free and premium training here](#).

If you have any questions, you can always email me - me@charlesfloate.co.uk

Thanks once again,

A handwritten signature in black ink that reads "Charles Floate". The signature is written in a cursive, flowing style with a long horizontal line extending from the end of the name.