The Ultimate Google Disavow Guide

Since its introduction in 2012 the Google Disavow Tool has become an indispensable instrument for website operators and online marketers It is the only method of positively mitigating backlink risks and maintaining off-page signals. If used correctly, the Disavow Tool becomes a one-way communication channel with Google. The wider application of the Disavow Tool and the vast amounts of data that it continues to generate for Google remains a source of speculation. It’s main purpose, at least from the website operator’s point of view, remains the dissociation of their website from PageRank passing backlinks that may otherwise hold organic search visibility down. When used carelessly, however, the Disavow Tool can spell doom for Google Rankings.

This Ultimate Google Disavow Guide, strongly influenced by the author’s professional experience while working for Google Search, both penalizing offending websites as well as lifting Google penalties, is an attempt to clarify frequent misbeliefs around the Google Disavow Tool and its application. At the same time, it is an effort to address commonly associated questions and provide a blueprint for anyone considering using the Disavow Tool in order to improve their website signals.

Disavow or not to disavow?

Ostensibly there is a central question: **does every website require a disavow file?** The unequivocal answer to that question is a resounding: no! Not every website needs a compelling disavow file. In fact, in the scale of things, relatively few websites need to actively monitor and manage

their backlink signals. Websites such as personal blogs, government platforms, charities, non-governmental websites, even small niche or local webshops frequently aren't in acute need to disavow spam backlinks. The reason for this is that these websites rarely conduct link building. Many thrive on direct traffic and have neither the capacity nor the desire to improve their Google rankings. Often their target audience is aware of and familiar with their presence. Hence type-in traffic represents almost all of the traffic they enjoy. Consequently, they rarely if ever actively pursue PageRank, passing backlinks that Google, in turn, may frown upon. With relatively few backlinks in all, disavowing is virtually a non-issue.

It is an entirely different situation when looking at commercial websites, such as online shops, price comparison platforms, market places, media outlets, portals or major brands. Their overriding commercial intent makes them susceptible to optimization, which may or may not always have been Google Webmaster Guidelines compliant. Google remains adamant with regard to, in Google's mind, not merit based backlinks, hence managing backlink risks is for these websites a critical part of conducting online business. They need to use the Disavow Tool as a shield protecting their organic rankings.

There's also a situation where every **affected website must use the Disavow Tool** in their defense. That's when a Google manual spam action aka manual penalty due to link building is in place. As a general rule, any Google penalty should be removed as swiftly as possible. However, a penalty in relation to backlinks, in particular, must be addressed immediately, since it progressively impacts the website's position in Google Search Results.

Any website held back by a manual penalty due to spam backlinks must use the Disavow Tool.

When to disavow?

There are a number of important, yet only two main factors which must be considered if backlinks may constitute a liability for the website: **the volume of incoming backlinks and their quality**. The first indicator can frequently be gauged almost instantly, by looking at the total number of backlinks reported in trustworthy third party tools such as Majestic. For example, the minimum number of backlinks ever recorded pointing to example.com is the sum of fresh and historic total combined. While no single tool is capable of providing an exact figure, approximately 300 million backlinks is a substantial number that in case of a commercial website may warrant a review and updating the disavow file. Majestic, similarly to other great data gathering tools recommended later in this guide, is unlikely to detect all backlinks ever to be in existence. Like all other commercial tools, it may not identify private (blog) backlink networks, created specifically in order to avoid third party detection. PBNs are however a ludicrous concept from its inception. Link building is explicitly done for Google, backlinks must be detectable by Google and therefore always pose a clear liability to the website's rankings.

Google Search Console, while indispensable in the process, can't be considered as the ultimate data source because of the build in reporting limit, capping samples at 100.000 backlinks. That having said, when looking for a tangible threshold 100.000 backlinks may be taken as a rule of thumb. Fewer backlinks likely do not warrant the effort required for disavowing. More backlinks may, potentially.

The second main indicator -quality- is significantly less simple to even ascertain, let alone to accurately assess.

Backlink quality depends on the type of anchor text used, anchor text distribution, the quality of content surrounding backlink anchors, as well as where else the same page links to. In a nutshell, it can only be analyzed by experienced human experts armed with powerful, purpose built tools which help to expedite the process. No tools however can fully replace this labour intensive, detail oriented approach. Manual analysis and investigating backlink quality requires **crawling backlink data** in a first, critical step.

Short of going through the entire process, there's one additional indicator which can help to gauge how much of an acute risk a backlink profile may pose: **the anchor text distribution**. While there are no hard thresholds to observe, the ground rule is that the more the top ten anchor texts appear optimized for the specific products or services offered, the higher the probability that PageRank passing link building was conducted at some point in time. And that consequently legacy and fresh backlinks are more likely to pose a serious risk. Several tools offer insights in this regard, including Ahrefs and Majestic, with varying depth. Evaluating all but the top 10 anchor texts is however superfluous, since commercial anchors tend to surface to the top anyway.

Highly commercial top 10 anchor text distribution, as in this example, usually is a strong indicator of backlink liabilities.

What to disavow?

Any backlink analysis must begin with aggregating as much **relevant backlink data** as possible. Google Search Console backlink samples are a stepping stone towards that purpose. As mentioned previously, these are limited, effectively reducing

their informative value for websites with substantial backlink profiles of 10 million backlinks and more. While it is possible to boost the GSC sample output by adding a multitude of patterns next to domain property, such as with www. and without www., both https and http as well as a combination of all the above and possibly other subdomains or directories, these steps remain a work-around. Google's continuous insistence on the matter that GSC samples are sufficient for any eventuality is true, yet only for relatively small websites. Websites which over time accumulated substantial backlink profiles must not rely on GSC samples alone. A cost-effective, yet time-consuming option can be collecting backlink data samples from other search engines webmaster tools as well.

Similarly to Google Search Console, the Bing Webmaster Tool BETA feature offers a free of charge glimpse at the backlink data. On a side note, Bing also allows webmasters concerned with Bing rankings to disassociate their websites from undesirable backlinks, not unlike Google does. This Ultimate Google Disavow Guide however is solely focussed on Google best practices.

Other leading search engines webmaster tools, such as Yandex or Baidu can be useful to collect yet more backlink data over a period of time. While the former however seems to take a long time before amassing and displaying backlink data, the latter poses a formidable language challenge for many website operators.

As mentioned, none of the search engines provided data samples, mentioned alone or even combined, is complete enough when reliable data is acutely needed. In a critical situation it is of the greatest importance that other third party data samples are also taken into consideration in order to generate a sufficient data sample for actionable results. Here again ideally a multitude of different tools such as Ahrefs, Majestic, SEMrush, Ryte and LinkResearchTools should be used

in order to accumulate and verify as much backlink data as possible. However, this redundant approach comes at a cost. The few high-quality services available come at a price, which can become substantial, once individual backlink data exports exceed 100 million backlinks. That modest number also dispels any lingering doubts whether Google Search Console backlink samples possibly could be sufficient for a through backlink analysis. Google Search Console capped 100 thousand backlink examples represent 0.1% sample of a 100 million backlink profile. A number hardly sufficient to cleanse a website of problematic link building past.

When the potentially time consuming process of collecting data is finalized, the samples collected are deduped and filtered in order to expedite the following, necessarily manual review process. Zero impact backlinks e.g. pointing to landing pages which are excluded in the websites robots.txt file or backlinks bearing a rel=”nofollow” attribute can be immediately dropped from consideration. They are inconsequential from a spam risk assessment perspective.

Both the attributes rel=”sponsored” and rel=”ugc”, which at this point remain scarcely applied novelties do not require any differentiated approach. They are exclusive to Google and not recognized by any other search engine, which is why if they were to be used at all, they must be applied in tandem with the industry-wide standard .

Similarly to backlinks that have no impact, there are also websites which can legitimately cross link, even using the most commercial anchor texts imagined, because they share ownership. In other words a website operator or organisation is at liberty to cross link their websites without risking violating Google Webmaster Guidelines. Hence the entire **website/domain portfolio should be excluded** from further analysis.

Contrary to that situation, subjectively high authority or

respectable brand websites should not be whitelabeled in a similar fashion. There are a lot of misconceptions around authority, most of all highlighting an elusive DA or domain authority value. The term so often tossed around is in-fact **not relevant to Google**. Experience shows that presumably respectable websites or brands frequently are in violation with Google linking policies. A fact not lost on the Google team. Therefore backlinks originating from what may be considered respectable sources must not be evaluated in any other way.

Equally high risk backlinks that are about to be removed entirely, changed or nofollowed, still need to be included in the evaluation process. If deemed a threat, they all must be included in the disavow file, despite the fact they may have changed already. Google may not have crawled them anew yet. Consequently changed signals may not be reflected in Google data and algorithms either, which makes such backlinks a lingering threat. If and when such backlinks are recrawled depends on the websites individual crawl budget allocation and management. And it can take some time, especially in case of low quality sites.

When commencing the analysis the backlinks **intent**, why they came to existence, and not ownership or origins remains in the focus of the investigation. Google neither asks nor cares how exactly spam backlinks came into existence, who created them or when. Therefore it isn’t necessary to record additional information for later processing or documentation purposes.

With fresh backlink data at hand, high probability spam backlinks can be grouped together. While **every website’s backlink profile is different** and constantly changing, there are backlinks that can with 100% certainty be regarded as a liability and securely disavowed. Templated spam, including auto generated websites with none, scraped or gibberish content are firmly within this group and must be disavowed.

High Risk TLDs

No TLD can be considered 100% spam. None should be disavowed merely on the grounds of the linking sites TLDs. There are however some TLDs which tend to over proportionally attract spam. The Spamhaus Project 2020 Top 10 statistics are rather revealing in this regard, even if the list isn't comprehensively long. Backlinks originating from websites hosted on domains such *as .tk, .gq, .top, .ml* or *.loan* to name just a few among many TLDs that can not be outright dismissed as spam, but in the backlink review process they can be filtered and grouped together. Often repetitive patterns in their **domain naming or URL structure, identical templates and gibberish, easy to spot content** help to make a swift risk assessment.

The situation is similar, yet slightly more nuanced, with spam free-hosting services. Every free-host does of course include some low quality websites. Some however notoriously fail to rid themselves of auto-generated spam, which in the past triggered a collective punishment reaction on Google's part. These arguably few publicly noticed instances demonstrate that Google did and does care about free-hosts. There is no need to preemptively include free-host services into the disavow file, however when dodgy free-hosted websites are identified in the course of the backlink analysis, there is no need to proceed with caution. Free-hosted spam backlinks justify including all of the service using the domain: operator to the disavow file.

Several TLDs are prone to attract templated spam websites, which must be disavowed.

Expired domains backlinks

While investigating backlink risk levels, some sites or, in this instance, domains are more easily recognized as spam and therefore harmful then others. Expired domains, that’s previously legitimate websites, dropped by their original operators just to be revived with scraped or templated content in the hope to benefit from reputation built in the past. They are a clear violation of Google Webmaster Guidelines and a **black hat SEO smoking gun**. Consequently all expired domains must be disavowed. Such sites are almost universally auto generated and therefore easily to spot and to filter. Where there is even a shadow of a doubt, the Internet Archive provides invaluable and free of charge services, showing most sites past records.

An example of a legitimate website before the domain expired.

And the same site after the domain expired. Such websites are generally to disavow.

Hacked sites backlinks

Similarly to expired domains, all legitimate yet compromised e.g. hacked websites, unknowingly linking through injected code without the legitimate operators consent, must be disavowed. Since this method as a trend is in decline, typically even for very large backlink profiles there will be only a hand-full of hacked websites in their backlink profile. This is the only group which may be revisited periodically to reassess the situation. Sites that have been cleansed of injected backlinks and content can be safely removed for the disavow file, however this is an optional step.

Press release backlinks

PageRank passing **press release backlinks, especially the ones bearing commercial anchor texts** as mentioned specifically in Google Linking Guidelines, must all be included in a compelling disavow file. Google has time and again highlighted their stands on press release link building and maintains their position that doing so is a clear Webmaster Guidelines offence. Veteran Google employees including John Muller have repeatedly reiterated on the specific point.

Affiliate backlinks

Google does not look unkindly at affiliate websites in general. PageRank passing affiliate backlinks however, are considered a thorn in the side, since they are not merit based as far as Google is concerned. Consequently, when managing backlink risks, all affiliate backlinks must be included in the disavow file. Coupon or special deals websites, as well as price comparison platforms, are worthy of special attention. By far not all of these legitimate services choose to ignore Google Linking Guidelines intentionally, however the ultimate responsibility to check, remains as with all backlinks with the site operator concerned for his or her websites Google rankings.

Directories backlinks

All SEO and link directories must be disavowed. At this point there are no legitimate reasons to make any exceptions. There are countless giveaways betraying the sole purpose for directories existence, which is passing PageRank. Among these are the fact that almost no directories are moderated, they lack topicality or any oversight and frequently even the domains used highlight SEO or links rather than any other editorial value. While frequently and rightly seen as a legacy

issue, directory backlinks remain a liability for websites even many years later.

Directories have lost credibility and should be disavowed.

Off-topic forums backlinks

As previously mentioned, PageRank passing spam backlinks have no expiration date. As long as they exist, they continue to pose a threat. Which is why even truly antiquated black hat link building methods like SEO directories and off-topic forum spamming must be included in the analysis. Especially the latter bares some serious legacy potential to cause harm, with efficient software solutions like XRumer on the market for a decade. Of course, not all forum backlinks are harmful to the target website. On the contrary, relevant, community-driven and moderated forum references can be great. And most importantly, drive converting traffic too. However, these are in a manual review easily separated from the spam entries, from junior form members that have little substance and no standing in the respective community. The latter type of spam backlinks must be disavowed.

Paid blogs backlinks

While the above mentioned types of link spam are usually swiftly identified based on similarities, it is paid blog posts that require more scrutiny. The decisive factor for determining the risk that paid blog posts cumulatively pose is intent. In other words, if the quality of the website linking, the depth of its contents, its unrestrictive linking policies, as well as the type of anchor text used and the landing page the backlinks point to indicate intent to benefit from passing PageRank, such backlinks pose a risk. They must be disavowed.

While paid blogging is rarely done with quality in mind, short of relying on tested CMS' and reliable, cheap hosting such as WordPress paid blogs do not necessarily display the same level of uniformity that allows for efficient filtering in the review process. Like all of the types of spam specifically mentioned in this guide, paid blog **backlinks require a manual review** in preparation for a risk assessment and in order to build a compelling disavow file.

When in-doubt these are some questions that may help ascertain, whether a backlink is a risk contributing factor and possibly should be disavowed:

- Is the backlink passing PageRank?

If it isn't, either because of a rel="nofollow" attribute or because the landing page is robots.txt-ed out, it does not pose a problem. All subsequent questions become obsolete. If the backlink is passing PageRank, addressing the following questions can help:

- Was the link paid for?
- Do you create and control the backlink?
- Was the backlink auto generated?

If any of these questions are met with a positive response, the backlink likely is a risk factor.

Ultimately, the key question that can help to determine if a backlink is legitimate as far as Google is concerned or not as much is: **Would you be comfortable sharing this link with a competitor or a Google Search employee?** That latter question answered rarely fails to sort the wheat from the chaff.

How to disavow?

Once the analysis is complete a few, albeit important steps remain. Google does provide some guidance regarding the formatting and disavow file limitations set. The two most

obvious limits, which are maximum 2MB file size and a total of 100.000 URLs constrain few, but the most heavily link spammed websites on the web. Avoiding individual, granular patterns and **always utilizing the domain: operator** on site level [e.g. domain:example.com] are basic best practices to abide. The correct file format of the .txt document which must be UTF-8 or 7-bit ASCII remains the last important point to consider. As long as it ends on .txt the file name used is inconsequential.

Website operators can verify the validity of their disavow file using the only free of charge Disavow File Testing Tool website, created by ex-Google engineer and fellow Search Engine Land author Fili Wiese. No tool, however, can detect contextual errors, such as disavowing one's own domains, which can and should legitimately cross-link. This is why a final check should ensure no relevant, legitimately linking domains that belong to the same website operator are included.

Before submission a new disavow file should be tested.
Another hack in the final step is to submit the disavow file individually to all verified Google Search Console patterns, including all combinations, such as with www. and without www., https and http as well as any other subdomains and directories. While Google officially recommends to focus exclusively on the primary property, this seemingly redundant effort, can help to protect the website from undesirable backlinks in a rare case of Google systems failing, whether temporarily or critically.

A finalized and double-checked disavow file is **best submitted without delay**. Backlink profiles of all websites evolve and change constantly. New backlinks come into existence, desirable ones as well as spam backlinks. That's why disavow files degenerate over time. That process happens faster for

websites with already relatively large backlink profiles. Which is why a finalized disavow file represents a temporary remedy. The value of which declines as time passes.

An updated or new disavow file is best submitted individually to all verified GSC patterns.

What’s next?

For Google an individual disavow file represents a website operator recommendation, not a directive. Whether it chooses to follow that recommandation in full or partially isn’t however disclosed via Google Search Console. For the website operator, the very same disavow file represents a **temporary protection shield**. For how long it may provide some level of confidence, depends on subsequent backlink growth and its quality. While there are few general rules to follow, a disavow file should be revisited and updated based on fresh data at least once per year, as part of an annual maintenance cycle. When that reiteration happens, previously disavowed and newly detected spam backlink patterns must be **combined into one new disavow file** before uploading. Merely uploading new patterns will inevitably and irreversibly delete previously submitted spam backlink patterns and thereby undo the past good work.

The submitted disavow file has at the same time **no impact on converting traffic** forthcoming from included backlinks. In other words, while backlink risks are mitigated, whatever traffic may originate from the same suspicious backlinks, isn’t affected.

Updating a disavow file can lead to a traffic growth spurt.

There are no methods of confidently predicting traffic trajectory after disavowing spam backlinks in Google. Three scenarios are distinctly possible. Websites traffic can stagnate, sharply increase as back link ballast is removed or drop significantly. Disavowing questionable backlinks isn't about growing traffic though. That's what an on- and off-page technical audit is for. Disavowing helps to maintain the website's off-page signal input. At the same time, the very fact that Google penalizes websites for building PageRank passing links demonstrates without any doubt that link building, including risk-taking link building in violation with Google Webmaster Guidelines, can work. Above and beyond, passing PageRank links as a ranking factor is here to stay and must be considered as an important SEO signal, especially for content discovery, searchbot crawl prioritization, user navigation and converting traffic. There are risk-averse, conversion-oriented alternative approaches towards building backlinks, as comprehensively covered in the How to build links article, which offers a new perspective. One that, if fully embraced, can also help to reduce the need to disavow backlinks in the future to come.

Opinions expressed in this article are those of the guest author and not necessarily Search Engine Land. Staff authors are listed here.

About The Author

Kaspar Szymanski is a founding member of Search Brothers and well-known search expert specializing in recovering websites from Google penalties and helping websites improve their rankings with SEO Consulting. Before founding SearchBrothers.com, Kaspar was part of the Google Search Quality team where he was a driving force behind global web spam tackling initiatives. He is the author of the ultimate guide to Google penalties and part of the Ask the SMXperts series.
http://feeds.searchengineland.com/~r/searchengineland/~3/wFRVgTBbDZU/the-ultimate-google-disavow-guide-334939