

Of Pirates and PageRank

by [Andrew McDiarmid](#) [1]
August 15, 2012

Google's latest change to its closely guarded search ranking algorithm has elicited responses ranging from raised eyebrows to guarded praise. The [announcement](#) [2] was short on details, but the upshot is that certain sites that Google's DMCA takedown data suggest are more likely to be offering copyright-infringing content will be demoted in search results, what some commentators have called a "[Pirate Penalty](#) [3]".

It's too early to fully judge the new initiative, but a truly voluntary measure should not be dismissed out of hand. It is certainly preferable to – and may indeed have a more meaningful impact on online infringement than – misguided blocking legislation like SOPA and PIPA. If the change results in steering more traffic to licensed sources and away from infringing sites, it could be a win-win: reducing infringement, encouraging the growth of lawful alternatives, and relieving some of the pressure for overbroad legislative "solutions" to the piracy problem.

It is a significant change, however, and is not without risks. As we [wrote](#) [4] to the Obama administration last week, voluntary measures to address online infringement have some advantages – chiefly their flexibility – but can lack the transparency and commitment to due process that we expect to be baked in to judicial and legislative actions. Of course, Google is free to write its algorithms any way it wants. However, Google is a powerful intermediary that can exert a major influence on a site's traffic, and as a matter of best practice the company should be upfront about what it is doing, and provide meaningful ways to correct mistakes and address unintended consequences.

Risk of Gaming and Abuse

One obvious risk is third parties gaming the system to disadvantage competitors. Gaming search results is nothing new to Google, but this change may mark a new and significant opportunity for third parties to push others down the ranks. What if someone decides to hamper a competitor's site by sending a flood of notices to Google to remove links to the competitor's site? What if, as often happens (see the cases against [YouTube](#) [5] and [Veoh](#) [6], e.g.), rightsholders are simply unhappy with a user-generated content site that itself complies with DMCA safe harbors, and nonetheless decide to route their notices to Google Search instead? In either scenario, a fully lawful site could turn up lower in Google's rankings. As Public Knowledge [has noted](#) [7], Google insists only "valid" notices will factor in under the new system, but valid in this case only means compliant with the formal requirements of the DMCA. At the volume it currently receives notices for links in Google Search ([around 5 million per month](#) [8]), the company has no way of verifying the claims made in every valid notice.

Transparency and close scrutiny of which sites are being affected and how will therefore be essential. Without compromising its need for secrecy where the algorithm is concerned, Google should make it clear what criteria it uses (a [recent report](#) [9] indicated that it is unsurprisingly more complicated than just the number of complaints received), how it is evaluating the change's effectiveness, and what safeguards are in place to prevent and correct for abuse so that legitimate sites aren't harmed.

Equally important for combating overbreadth is having a process for determining if and when mistakes are being made. The data at Google's disposal – DMCA notices to remove links to particular pages from its search index and its crawler's snapshot of a given site – are imprecise indicators. Lawful sites (not to mention lawful content present at largely infringing sites) will undoubtedly be affected by the change. It is in Google's interest to reduce and mitigate mistakes, but this requires data about what mistakes are being made. There must therefore be some recourse available to site owners who believe they have been wronged.

Since Google does not necessarily have a direct relationship with the sites in its index, it may not have a way to notify owners of "demoted" sites. Thus a key data source and safeguard against abuse in the DMCA, the counter-notice provision, may be unavailable. To its credit, Google has pledged to notify anyone who registers their site with its Webmaster Tools about removal requests for their site and to honor counter-notices. The announcement of the algorithm change did not, however, note this as an option. In future communications and FAQs, and as part of Webmaster Tools itself, Google should make it clear that third-party complaints can affect rankings, share complaint data with affected site owners, and provide information about the recourse available in case of mistakes.

Keep it Voluntary

Lastly, Google is just one company and not Congress or a court of law. Nothing in Google's adoption of this new signal for ranking should imply that it is setting a new standard or prerequisite for DMCA safe-harbor or for successfully defending against claims of secondary liability. Smaller sites may lack the resources and data to follow suit, and it would be inappropriate to turn one voluntary action by one major search engine into a de facto mandate for all. Moreover, any effort to turn voluntary efforts into mandates may chill companies from good-faith experimentation with novel approaches to fighting infringement.

Nor should Google's decision to engage in voluntary downranking of possibly infringing sites be read as conclusive evidence that those sites are hubs for infringement. Only courts can determine definitively what is infringing and what is not, and it would be unfair to impute guilt based on one search engine's subjective assessment of a site's quality.

Since before the SOPA and PIPA fight, Google has been striving to position itself as a good actor on copyright issues, and it deserves some credit for stepping up and putting its algorithm where its mouth is. Given the power of its rankings on the web, however, Google must tread carefully and transparently, keeping a close eye on the impact of the algorithm change and giving regular updates that allow others to do the same.

Ultimately, the success of Google's move here will be determined by the results that rise to the top of the rankings. If the downranked results are simply replaced by different infringers – or the downranking is undone by affected sites' efforts to work around their demotion – then this change will amount to little more than a speed bump. The availability of compelling and convenient lawful offerings to meet consumer demand for content is a critical part of stopping online infringement, and Google can't make people click on those links if they're not there.

- [search engine](#)
- [page rank](#)
- [Intellectual property](#)
- [Google](#)
- [infringement](#)
- [copyright](#)

The copyright © 2013 by the Center for Democracy & Technology. CDT can be freely copied and used as long as you make no substantive changes and clearly give us credit. [Details.](#)

Source URL: <https://cdt.org/blogs/andrew-mcdiarmid/1508pirates-and-pagerank>

Links:

[1] <https://cdt.org/personnel/andrew-mcdiarmid>

[2] <http://insidesearch.blogspot.com/2012/08/an-update-to-our-search-algorithms.html>

- [3] http://searchengineland.com/google-many-popular-sites-will-escape-pirate-penalty-130289?utm_source=feedburner&utm_medium=feed&utm_campaign=feed-main
- [4] <https://www.cdt.org/blogs/david-sohn/0908cdt-weighs-copyright-enforcement-strategy>
- [5] <https://www.cdt.org/blogs/david-sohn/0504court-ruling-youtube-appeal-mostly-win-internet>
- [6] <https://www.eff.org/cases/umg-v-veoh>
- [7] <http://www.publicknowledge.org/blog/google-unintended>
- [8] <http://www.google.com/transparencyreport/removals/copyright/>
- [9] <http://searchengineland.com/google-many-popular-sites-will-escape-pirate-penalty-130289>