A Complete Guide to the Google Penguin Algorithm Update by @TaylorDanRW

Editor's note: *This post is part of an ongoing series looking back at the history of Google algorithm updates. Enjoy!*

In 2012, Google officially launched the "webspam algorithm update," an algorithm specifically targeting link spam and manipulative link building practices.

The webspam algorithm later became known (officially) as the Penguin algorithm update via a tweet from Matt Cutts, who was then head of the Google webspam team. While Google officially named the algorithm Penguin, there is no official word on where this name came from.

ADVERTISEMENT

The Panda algorithm name came from one of the key engineers involved with it, and it's more than likely that Penguin originated from a similar source. One of my favorite Penguin naming theories is that it pays homage to The Penguin, from DC's Batman.

> *Minor weather report: We pushed 1st Penguin algo data refresh an hour ago. Affects <0.1% of English searches. Context: http://t.co/ztJiMGMi*
>
> *— Matt Cutts (@mattcutts) May 26, 2012*

Prior to the Penguin algorithm, link volume played a larger part in determining a webpage's scoring when crawled, indexed, and analyzed by Google.

This meant when it came to ranking websites by these scores for search results pages, some low-quality websites and pieces of content appeared in more prominent positions of the organic search results than they should have.

Why Google Penguin Was Needed

Google's war on low-quality started with the Panda algorithm, and Penguin was an extension and addition to the arsenal to fight this war.

Penguin was Google's response to the increasing practice of manipulating search results (and rankings) through black hat link building techniques. Cutts, speaking at the SMX Advanced 2012 conference, said:

> *We look at it something designed to tackle low-quality content. It started out with Panda, and then we noticed that there was still a lot of spam and Penguin was designed to tackle that.*

The algorithm's objective was to gain greater control over, and reduce the effectiveness of, a number of black hat spamming techniques.

By better understanding and process the types of links websites and webmasters were earning, Penguin worked toward ensuring that natural, authoritative and relevant links rewarded the websites they pointed to, while manipulative and spammy links were downgraded.

Initial Launch & Impact

When Penguin first launched in April 2012, it affected more than 3 percent of search results, according to Google's own estimations.

> *Minor weather report: We pushed 1st Penguin algo data refresh an hour ago. Affects <0.1% of English searches. Context: http://t.co/ztJiMGMi*
>
> *— Scary Matt Cutts (@mattcutts) May 26, 2012*

Penguin 2.0, the fourth update (including the initial launch) to the algorithm was released in May 2013, and affected roughly 2.3 percent of all queries.

Key Google Penguin Updates & Refreshes

There have been a number of updates and refreshes to the Penguin algorithm since it was launched in 2012, and possibly a number of other tweaks that have gone down in history as unknown algorithm updates.

Penguin 1.1: March 26, 2012

This wasn't a change to the algorithm itself, but a refresh of the data within it.

In this instance, websites that had initially been affected by the launch who had been proactive in clearing up their link profiles saw some recovery, while others who hadn't been caught by Penguin first time round saw an impact.

This was the first data refresh.

Penguin 1.2: October 5, 2012

While this was another data refresh, I feel this is worth mentioning as it didn't only affect queries in the English language, but also affected international queries.

> *Weather report: Penguin data refresh coming today. 0.3% of English queries noticeably affected. Details:*

> *http://t.co/Esbi2ilX*
>
> *— Scary Matt Cutts (@mattcutts) October 5, 2012*

Penguin 2.0: May 22, 2013

This was a more technically advanced version of the Penguin algorithm and changed how the algorithm impacted search results.

Penguin 2.0 impacted around 2.3 percent of English queries, as well as other languages proportionately.

This was also the first Penguin update to look deeper than the websites homepage and top-level category pages for evidence of link spam being directed to the website.

The first refresh to Penguin 2.0 (2.1) came on October 24 of the same year. It affected a further 1 percent of queries.

While there was no official explanation from Google, data suggests that the 2.1 data refresh also advanced on how deep Penguin looked into a website and crawled deeper and conducted further analysis as to whether spammy links were contained.

Penguin 3.0: October 17, 2014

While this was named like a major update, it was, in fact, another data refresh; allowing those impacted by previous updates to emerge and recover, while many others who had continued to utilize spammy link practices, and had escaped the radar of the previous impacts saw an impact.

Googler Pierre Far confirmed this through a post on his Google+ profile and that the update would take a “few weeks” to roll out fully.

Far also stated that this update affected less than 1 percent of English search queries.

Penguin 4.0, September 23, 2016

Almost two years after the 3.0 refresh, the final Penguin algorithm update was launched.

The biggest change with this iteration was that Penguin became a part of the core algorithm.

When algorithm transcends to become a part of the core, it doesn’t mean that the algorithm’s functionality has changed or may change dramatically again. It means that Google’s perception of the algorithm has changed, not the algorithm itself.

Now running concurrently with the core, Penguin evaluates websites and links in real-time. This meant that you can see (reasonably) instant impacts of your link building or remediation work.

The new Penguin also wasn’t closed-fisted in handing out link-based penalties but rather devalued the links themselves. This is a contrast to the previous Penguin iterations, where the negative was punished.

That being said, studies and, from personal experience, algorithmic penalties relating to backlinks still do exist.

Data released by SEO professionals (e.g., Michael Cottam), as well as seeing algorithmic downgrades lifted through disavow files after Penguin 4.0, enforce this belief.

Penguin Algorithmic Downgrades

Soon after the Penguin algorithm was introduced, webmasters and brands who had used manipulative link building techniques or filled their backlink profiles with copious amounts of low-quality links began to see decreases in their organic traffic and rankings.

Not all Penguin downgrades were site-wide – some were partial and only affected certain keyword groups that had been heavily spammed and over optimized, such as key products and in some cases even brand.



A website impacted by a Penguin penalty, which took 17 months to lift.

The impact of Penguin can also pass between domains, so changing domain and redirecting the old one to the new can cause more problems in the long run.

Experiments and research shows that using a 301 or 302 redirect won't remove the effect of Penguin, and in the Google Webmasters Forum, John Mueller confirmed that using a meta refresh from one domain to a new domain could also cause complications.

> *In general, we recommend not using meta-refresh type redirects, as this can cause confusion with users (and search engine crawlers, who might mistake that for an attempted redirect).*

Google Penguin Recovery

The disavow tool has been an asset to SEO practitioners, and this hasn't changed even now that Penguin exists as part of the core algorithm.

As you would expect, there have been studies and theories published that disavowing links doesn't, in fact, do anything to help with link-based algorithmic downgrades and manual actions, but this has theory has been shot down by Google representatives publicly.

That being said, Google recommends that the disavow tool

should only be used as a last resort when dealing with link spam, as disavowing a link is a lot easier (and a quicker process in terms of its effect) than submitting reconsideration requests for good links.

Monitoring backlinks is also an essential task, as sometimes the industry we work in isn't entirely honest and negative SEO attacks can happen. This also means that using the disavow feature without a clear sign of an algorithmic penalty or a notification of a manual action.

Interestingly, however, a poll conducted by SEJ in September found that 38 percent of SEOs never disavow backlinks. Going through a backlink profile, and scrutinizing each linking domain as to whether it's a link you want or not, is not a light task.

Google recommends that you attempt to outreach to websites and webmasters where the bad links are originating from first and request their removal before you start disavowing.

While this is probably the most effective way to recover from a link-based penalty, it isn't always necessary. The Penguin algorithm also takes into account the link profile as a whole, and the volume of high-quality, natural links versus the number of spammy links.

While in the instances of a partial penalty (impacting over-optimized keywords) the algorithm may still affect you, the essentials of backlink maintenance and monitoring should keep you covered.

Some webmasters even go as far as including "terms" within the terms and conditions of their website and actively outreaching to websites they don't feel should be linking to them:

Website terms and conditions regarding linking to the website in question.

No Recovery in Sight?

Sometimes after webmasters have gone to great lengths to clean up their link profiles, and even after a known Penguin refresh, they still don't see an increase in traffic or rankings.

There are a number possible reasons behind this, including:

- The initial traffic and ranking boost seen prior to the algorithmic penalty was unjustified (and likely short-term) and came from the bad backlinks.
- When links have been removed, no efforts have been made to gain new backlinks of greater value.
- Not all the negative backlinks have been disavowed/a high enough proportion of the negative backlinks have been removed.
- The issue wasn't link-based, to begin with.

Penguin Myths & Misconceptions

One of the great things about the SEO industry and those involved in it is that it's a very active and vibrant community and there are always new theories and experiment findings being published online daily.

Naturally, this has led to a number of myths and misconceptions being born about Google's algorithms. Penguin is no different.

Here are a few myths and misconceptions about the Penguin algorithm we've seen over the years.

Myth: Penguin Is a Penalty

One of the biggest myths about the Penguin algorithm is that people call it a penalty (or what Google refers to as a manual action).

Despite the fact that an algorithmic change and a penalty can both cause a big downturn in website rankings, there are some pretty drastic differences between them.

A penalty (or manual action) happens when a member of Google's webspam team has responded to a flag, investigated and felt the need to enforce a penalty on the domain. You will receive a notification through Google Search Console relating to this manual action.

When you get hit by a manual action, not only do you need to review your backlinks and submit a disavow for the spammy ones that go against Google's guidelines, but you also need to submit a reconsideration request to the Google webspam team.

If successful, the penalty will be revoked, and if unsuccessful it's back to reviewing the backlink profile.

A Penguin downgrade happens without any involvement of a Google team member. It's all done algorithmically.

Previously, you would have to wait for a refresh or algorithm update, but now Penguin runs in real time so recoveries can happen a lot faster (if enough remediation work has been done).

Myth: Google Will Notify You if Penguin Hits Your Site

Another myth about the Google Penguin algorithm is that you will be notified if it has been applied.

Unfortunately, this isn't true. The Search Console won't notify you that your rankings have taken a dip because of the application of the Penguin.

Again, this shows the difference between an algorithm and a penalty – you would be notified if you were hit by a penalty. However, the process of recovering from Penguin is remarkably

similar to that of recovering from a penalty.

Myth: Disavowing Bad Links Is the Only Way to Reverse a Penguin Hit

While this tactic will remove a lot of the low-quality links, it is utterly time-consuming and a potential waste of resources.

Google Penguin looks at the percentage of good quality links compared to those of a spammy nature.

So, rather than focusing on manually removing those low-quality links, it may be worth focusing on increasing the number of quality links your website has. This will have a better impact on the percentage Penguin takes into account.

Myth: You Can’t Recover From Penguin

Yes, you can recover from Penguin.

It is possible, but it will require some experience in dealing with the fickle nature of Google algorithms.

The best way to shake off the negative effects of Penguin is to forget all of the existing links on your website, and begin to gain original editorially-given links.

The more of these quality links you gain, the easier it will be to release your website from the grip of Penguin.

Image Credits
Featured Image: Shutterstock, modified by Danny Goodwin
Screenshots taken by author