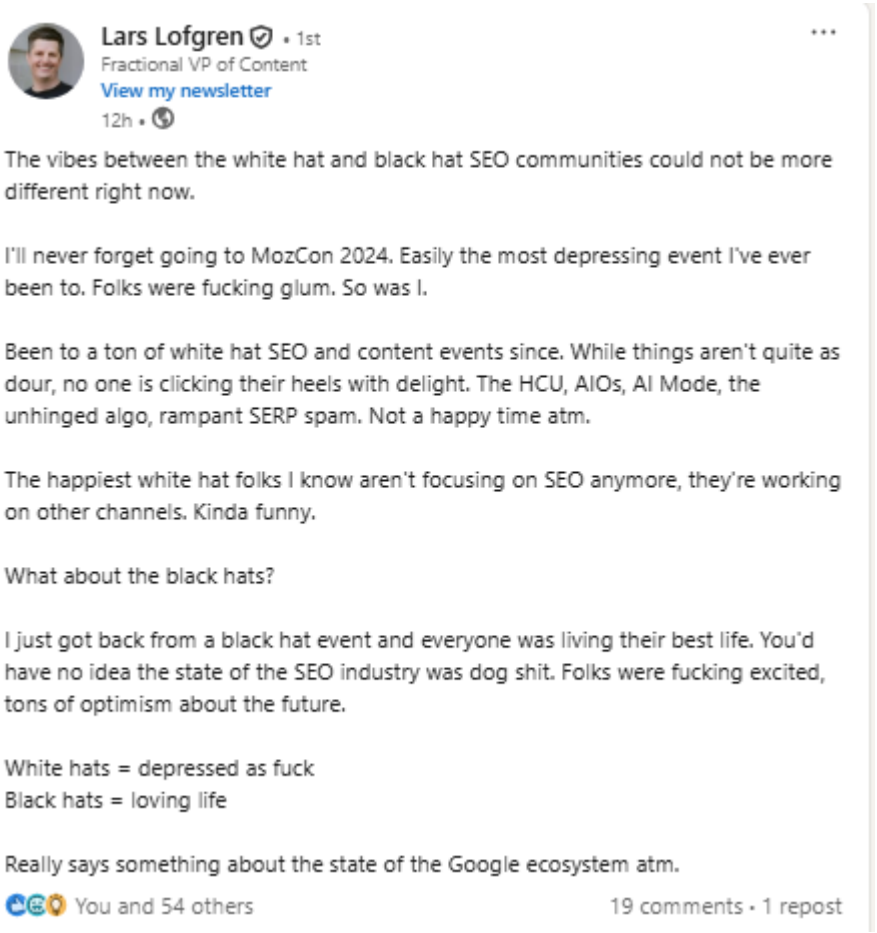
Why Google's Spam Problem Is Getting Worse

Spam is back in search. And in a big way.

Honestly, I don't think Google can handle this at all. The scale is unprecedented. They went after publishers manually with the site reputation abuse update. More expired domain abuse is reaching the top of the SERPs than at any time I can remember in recent history. They're fighting a losing battle, and they've taken their eye off the ball.



Lars Lofgren • 1st
Fractional VP of Content
View my newsletter
12h •

The vibes between the white hat and black hat SEO communities could not be more different right now.

I'll never forget going to MozCon 2024. Easily the most depressing event I've ever been to. Folks were fucking glum. So was I.

Been to a ton of white hat SEO and content events since. While things aren't quite as dour, no one is clicking their heels with delight. The HCU, AIOs, AI Mode, the unhinged algo, rampant SERP spam. Not a happy time atm.

The happiest white hat folks I know aren't focusing on SEO anymore, they're working on other channels. Kinda funny.

What about the black hats?

I just got back from a black hat event and everyone was living their best life. You'd have no idea the state of the SEO industry was dog shit. Folks were fucking excited, tons of optimism about the future.

White hats = depressed as fuck
Black hats = loving life

Really says something about the state of the Google ecosystem atm.

You and 54 others 19 comments • 1 repost

In a microcosm, this is what's happening (Image Credit: Harry Clarkson-Bennett)

A few years ago, search was getting on top of the various spam issues "creative" SEOs were trialling. The prospect of being nerfed by a spam update and Google's willingness to invest and care in the quality of search seemed to be winning the war.

Trying to recover from these penalties is nothing short of disastrous. Just ask anybody hit by the Helpful Content update.

But things have shifted. AI is haphazardly rewriting the rules, and big tech has bigger, more poisonous fish to fry. This is not a great time to be a white hat SEO.

TL;DR

1. Google is currently losing the war against spam, with unprecedented scale driven by AI-generated slop, and expired domain and PBN abuse.
2. Google's spam detection monitors four key groups of signals – content, links, reputational, and behavioral.
3. Data from the Google Leak suggests its most capable detection focuses on **link velocity and anchor text.**
4. AI "search" is dozens of times more expensive than traditional search. This enormous cost and focus on new AI products is leading to underinvestment in core spam-fighting.

How Does Google's Spam Detection System Work?

Via SpamBrain. Previously, the search giant rolled out Penguin, Panda, and RankBrain to make better decisions based on links and keywords.

And right now, badly.

SpamBrain is designed to identify content and websites engaging in spammy activities with apparently "shocking" accuracy. I don't know whether shocking in this sense is meant in a positive or negative way right now, but I can only parrot what is said.

Over time, the algorithm learns what is and isn't spam. Once it has clearly established signals associated with spammy sites, it's able to create a neural network.

Much like the concept of seed sites, if you have the spammiest websites mapped out, you can accurately score everyone else against them. Then you can analyse signals at scale – content, links, behavioral, and reputational signals – to group sites together.

- **Inputs** (content, linking reputational and behavioral signals).
- **Hidden layer** (clustering and comparing each site to known spam ones).
- **Outputs** (spam or not spam).

If your site is bucketed in the same group as obviously spammy sites when it comes to any of the above, that is not a good sign. The algorithm works on thresholds. I imagine you need to sail pretty close to the wind for long enough to get hit by a spam update.

But if your content is relatively thin and low value add, you're probably halfway there. Add some dangerous links into the mix, some poor business decisions (parasite SEO being the most obvious example), and scaled content abuse, and you're doomed.

What Type Of Spam Are We Talking About Here?

Google notes the most egregious activities here. We're talking:

- Cloaking.
- Doorway abuse.
- Expired domain abuse.
- Hacked content.

- Hidden text and content.
- Keyword stuffing.
- Link spam.
- Scaled content abuse.
- Site reputation abuse.
- Thin affiliate content.
- UGC spam.

Lots of these are grossly intertwined. Expired domain abuse and PBNs. Keyword stuffing is a little old hat, but link spam is still very much alive and well. Scaled content abuse is at an all-time high across the internet.

The more content you have spread across multiple, semantically similar websites, the more effective you can be. Using exact and partial match anchors to leverage your authority towards "money" pages, the richer you will become.

Let's dive into the big ones below.

Fake News

Google Discover – Google's engagement baiting, social network-lite platform – has been hit by the unscrupulous spammers in recent times. There have been several instances of fake, AI-driven content reaching the masses. It's become so prevalent, it has even reached legacy media sites (woohoo).



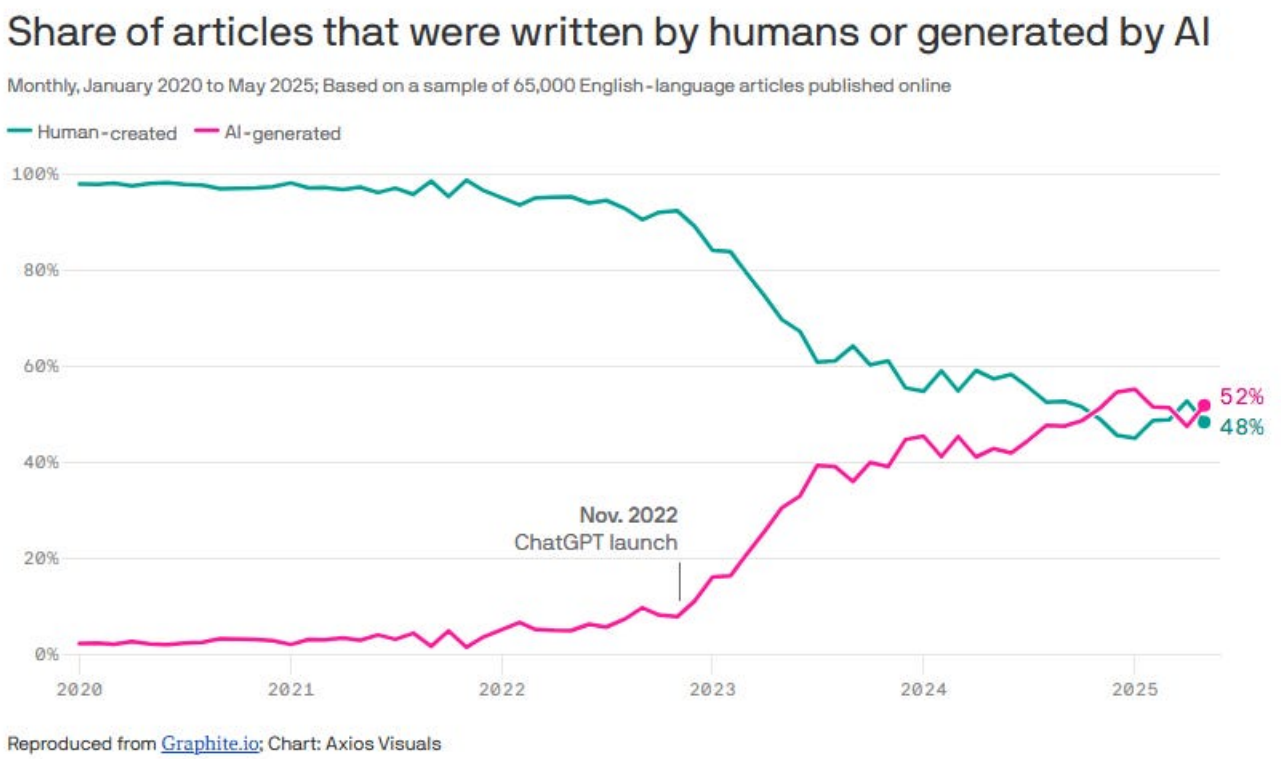
Millions of page views have been sent to expired and drop domain abusers (Image Credit: Harry Clarkson-Bennett)

From changing the state pension age to free bus passes and TV licenses, the spammers know the market. They know how to incite emotions. Hell hath no fury like a pensioner scorned, and while you can forgive the odd slip-up, nobody can be this generous.

The people who have been working by the book are being sidelined. But the opportunities in the black hat world are booming. Which is, in fairness, quite fun.

Scaled Content Abuse

At the time of writing, over 50% of the content on the internet is AI slop. Some say more. From nearly a million pages analyzed this year, Ahrefs says 74% contain AI-content. What we see is just what slips through the mammoth-sized cracks.



Not hard to see what the problem is… (Image Credit: Harry Clarkson-Bennett)

According to award-winning journalist Jean-Marc Manach's research, he has found over 8,300 AI-generated news websites in French and over 300 in English (the tip of the iceberg, trust me).

He estimates two of these site owners have become millionaires.

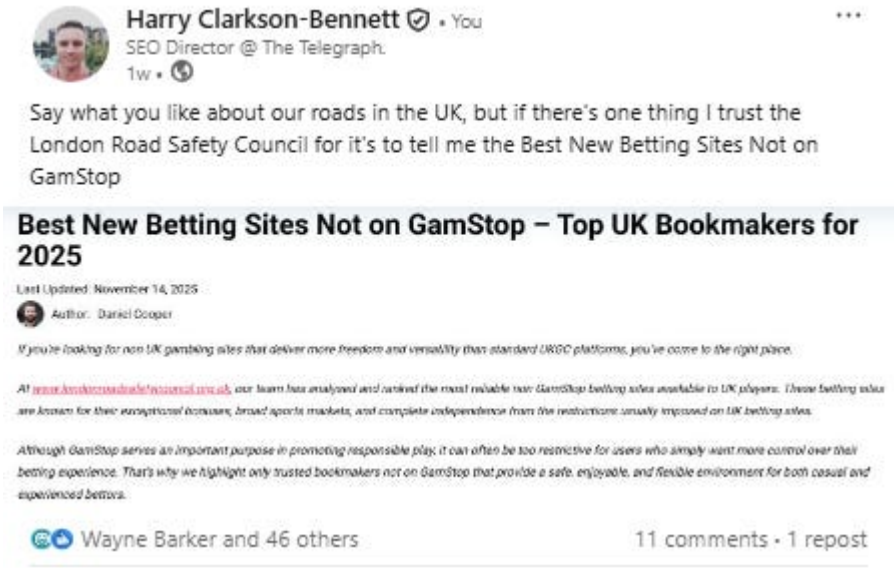
By leveraging authoritative, expired domains and PBNs (more on that next), SEOs – the people still ruining the internet – know how to game the system. By faking clicks, manipulating engagement signals, and utilizing past link equity effectively.

Expired Domain Abuse

The big daddy. Black hat ground zero.

If you engage even a little bit with a black hat community, you'll know how easy it is right now to leverage expired domains. In the example below, someone had bought the London

Road Safety website (a once highly authoritative domain) and turned it into a single-page “best betting sites not on GamStop” site.



This is just one example of many (Image Credit: Harry Clarkson-Bennett)

Betting and crypto are ground zero for all things black hat, just because there’s so much money involved.

I’m not an expert here, but I believe the process is as follows:

1. Purchase an expired, valuable domain with a strong, clean backlink history (no manual penalties). Ideally, a few of them.
2. Then you can begin to create your own PBN with unique hosting providers, nameservers, and IP addresses, with a variety of authoritative, aged, and newer domains.
3. This domain(s) then becomes your equity/authority stronghold.
4. Spin up multiple TLD variations of the domain, i.e., instead of .com it becomes .org.uk.
5. Add a mix of exact and partial match anchors from a PBN to the money site to signal its new focus.
6. Either add a 301 redirect for a short period of time to

the money variation of the domain or canonicalize to the variation.

These scams are always short-term plays. But they can be worth tens of hundreds of thousands of pounds when done well. And they are back, and I believe more valuable than ever.

> *Right now, I think it's as simple as buying an old charity domain, adding a quick reskin and voila. A 301 or equity passing tactic and your single page site about 'best casinos not on gamstop' is printing money. Even in the English speaking market.*
>
> *According to notorious black hat fella Charles Floate, some of these companies are laundering hundreds of thousands of pounds a month.*

PBNs

A PBN (or Private Blog Network) is a network of websites that someone controls that link back to the money site. The variation of the site designed to generate typically advertising or affiliate revenue.

A private blog network has to be completely unique from each other. They cannot share breadcrumbs that Google can trace. Each site needs a standalone:

- Hosting provider.
- IP address.
- Nameserver.

The reason PBNs are so valuable is you can build up an enormous amount of link equity and falsified topical authority to mitigate risk. Expired domains are risky because they're expensive, and once they get a penalty, they're doomed. PBNs spread the risk. Like the head of a Hydra, one dies; another rises up.

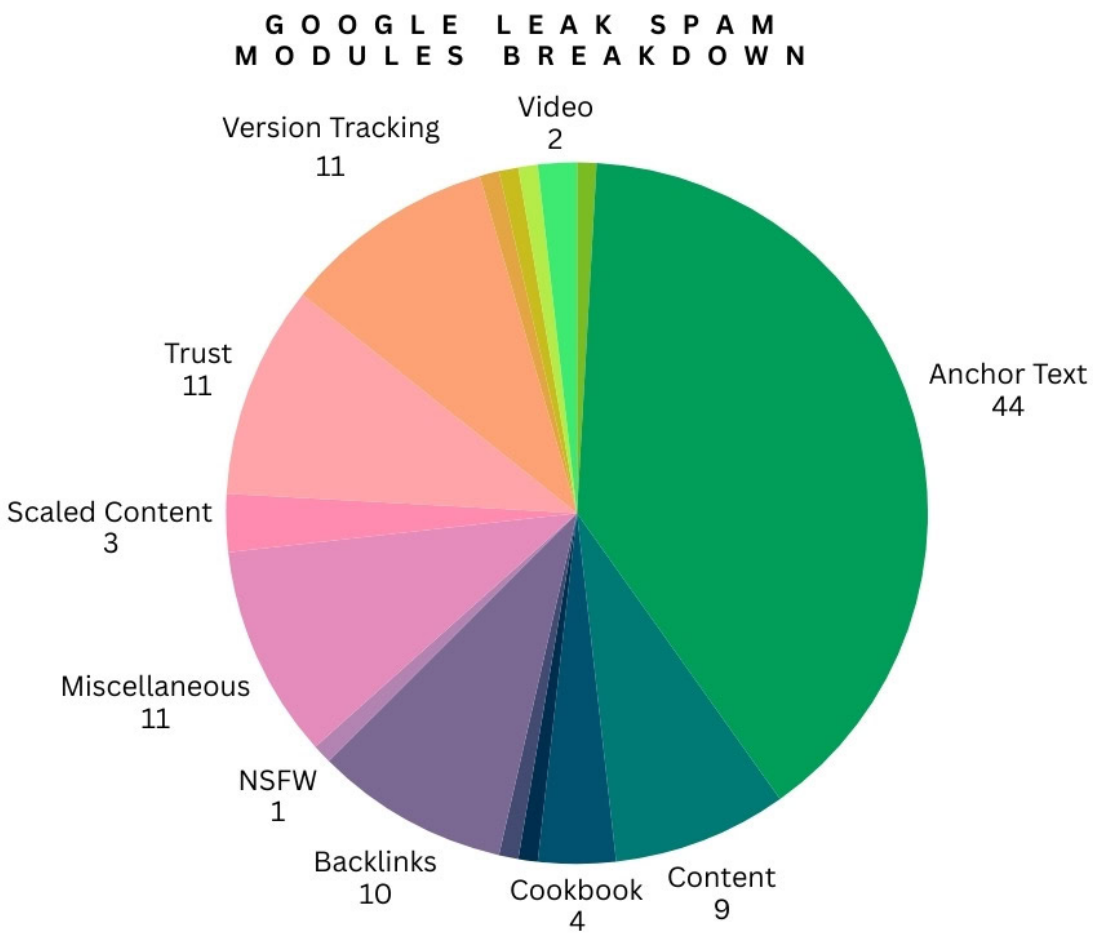
> *Protecting the tier 1 asset (the purchased aged or expired domain) is paramount. Instead of pointing links directly to the money site, you can link to the sites that link to the money site.*
>
> *This indirectly boosts the value of the money site, protecting it from Google's prying eyes.*

What Does The Google Leak Show About Spam?

As always, this is an inexact science. Barely even pseudo-science really. I've got the tinfoil hat on and a lot of string connecting wild snippets of information around the room to make this work. You should follow Shaun Anderson here.

If I take every mention of the word "spam" in the module names and descriptions, there are around 115, once I've removed any nonsense. Then we can categorize those into content, links, reputational, and behavioral signals.

Taking it one step further, these modules can be classified as relating to things like link building, anchor text, content quality, et al. This gives us a rough sense of what matters in terms of *scale*.



Anchor text makes up the lion's share of spammy modules based on data from the Google Leak (and my own flawed categorization)(Image Credit: Harry Clarkson-Bennett)

A few examples:

- **spambrainTotalDocSpamScore** calculates a document's overall spam score.
- **IndexingDocjoinerAnchorPhraseSpamInfo** and **IndexingDocjoinerAnchorSpamInfo modules** identify spammy anchor phrases by looking at the number, velocity, the days the links were discovered, and the time the spike ended.
- **GeostoreSourceTrustProto** helps evaluate the trustworthiness of a source.

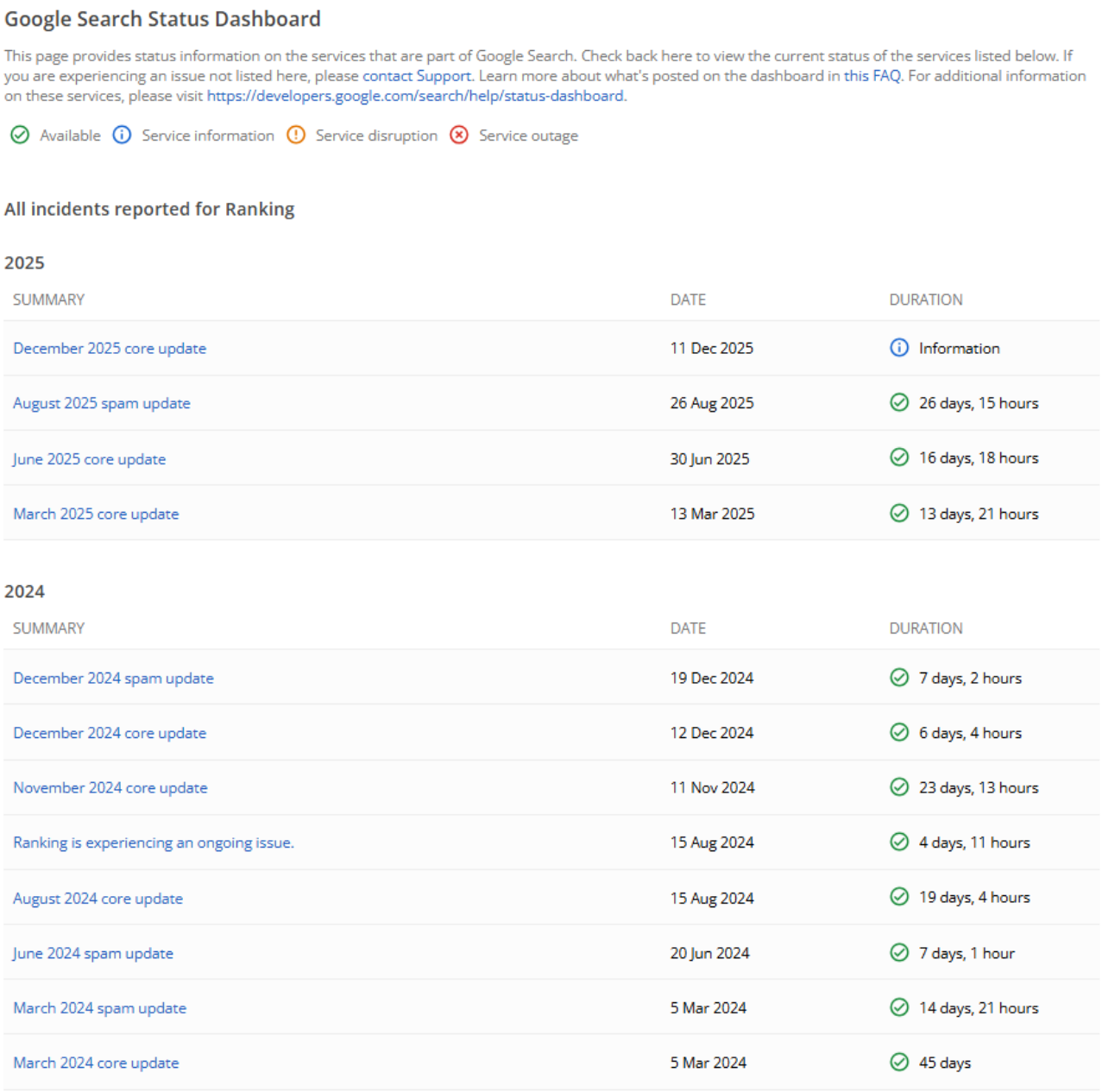
Really, the takeaway is how important links are from a spam sense. Particularly, **anchor text**. The velocity at which you gain links matters. As does the text and surrounding content. Linking seems to be where Google's algorithm is most capable of identifying red and amber flags.

If your link velocity graph spiked with exact match anchors to highly commercial pages, that’s a flag. Once a site is pinged for this type of content or link-related abuse, the behavioral and reputational signals are analysed as part of SpamBrain.

If these corroborate and your site exceeds certain thresholds, you’re doomed. It’s why this has (until recently) been a relatively fine art.

Ultimately, They’re Just Investing Less In Traditional Search

As Martin McGarry pointed out, they just care a bit less … They have bigger, more hallucinogenic fish to fry.



Google Search Status Dashboard

This page provides status information on the services that are part of Google Search. Check back here to view the current status of the services listed below. If you are experiencing an issue not listed here, please contact Support. Learn more about what's posted on the dashboard in this FAQ. For additional information on these services, please visit https://developers.google.com/search/help/status-dashboard.

Available | Service information | Service disruption | Service outage

All incidents reported for Ranking

2025

SUMMARY	DATE	DURATION
December 2025 core update	11 Dec 2025	Information
August 2025 spam update	26 Aug 2025	26 days, 15 hours
June 2025 core update	30 Jun 2025	16 days, 18 hours
March 2025 core update	13 Mar 2025	13 days, 21 hours

2024

SUMMARY	DATE	DURATION
December 2024 spam update	19 Dec 2024	7 days, 2 hours
December 2024 core update	12 Dec 2024	6 days, 4 hours
November 2024 core update	11 Nov 2024	23 days, 13 hours
Ranking is experiencing an ongoing issue.	15 Aug 2024	4 days, 11 hours
August 2024 core update	15 Aug 2024	19 days, 4 hours
June 2024 spam update	20 Jun 2024	7 days, 1 hour
March 2024 spam update	5 Mar 2024	14 days, 21 hours
March 2024 core update	5 Mar 2024	45 days

Image Credit: Harry Clarkson-Bennett

In 2025, we have had four updates, with a duration of c. 70 days. In 2024, we had seven that lasted almost 130 days. Productivity levels we can all aspire to.

It’s Not Hard To Guess Why…

The bleeding-edge search experience is changing. Google is rolling out preferred publisher sources globally and inline linking more effectively in its AI products. Much-needed changes.

I think we’re seeing the real-time moulding of the new search

experience in the form of The Google Web Guide. A personalized mix of trusted sources, AI Mode, a more classic search interface, and something inspirational. I suspect this might be a little like a Discover-lite feed. A place in the traditional search interface where content you will almost certainly like is fed to you to keep you engaged.

Unconfirmed, but apparently, Google has added persona-driven recommendation signals and a private publisher entity layer, among other things. Grouping users into cohorts is I believe a fundamental part of Discover. It's what allows content to go viral.

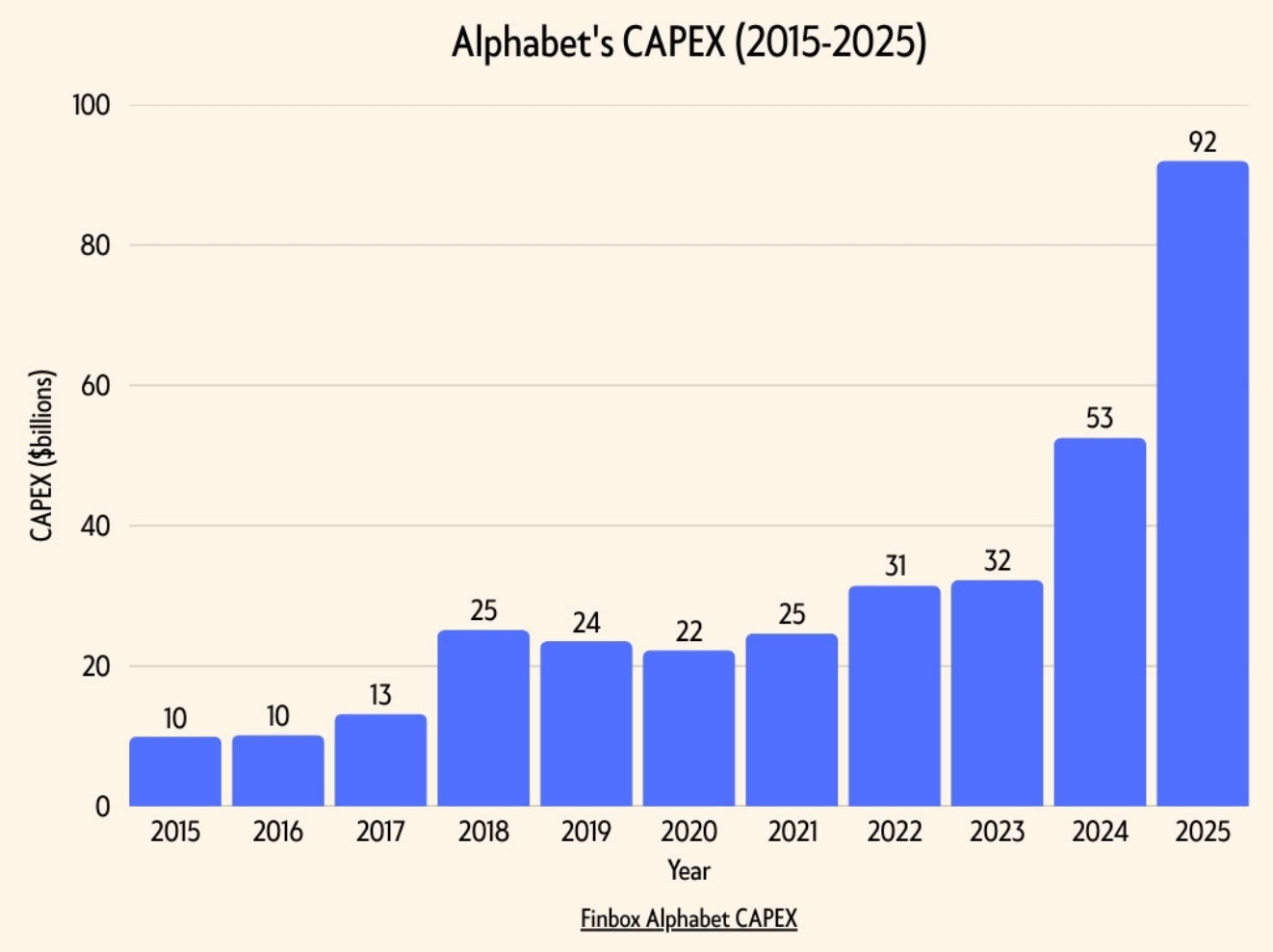
Once you understand enough about a user to bucket them into specific groups, you can saturate a market over the course of a few days Discover. Less even. But the problem is the economics of it all. Ten blue links are cheap. AI is not. At any level.

> *According to Google, when someone chooses a preferred source, they click through to that site twice as often on average. So I suspect it's worth taking seriously.*

Why Are AI Searches So Much More Expensive?

Google is going to spend $10 billion more this year than expected due to the growing demand for cloud services. YoY, Google's CAPEX spend is nearly double 2024's $52.5 billion.

It's not just Google. It's a Silicon Valley race to the bottom.



2025 has been extrapolated, but on course for $92 billion this year (Image Credit: Harry Clarkson-Bennett)

While Google hasn't released public information on this, it's no secret that AI searches are significantly more expensive than the classic 10 blue links. Traditional search is largely static and retrieval-based. It relies on pre-indexed pages to serve a list of links and is very cheap to run.

An AI Overview is generative. Google has to run a large language model to summarize and generate a natural language answer. AI Mode is significantly worse. The multi-turn, conversational interface processes the entire dialogue *in addition* to the new query.

Given the query fan-out technique – where dozens of searches are run in parallel – this process demands significantly more computational power.

Custom chips, efficiencies, and caching can reduce the cost of

this. But this is one of Google's biggest challenges. I suspect exactly why Barry believes AI Mode won't be the default search experience. I'd be surprised if it isn't just applied at a search/personalization level, too. There are plenty of branded and navigational searches where this would be an enormous waste of money.

And these guys really love money.

> *According to The IET, if the population of London (>9.7 million) asked ChatGPT to write a 100-word email this would require 4,874,000 litres of water to cool the servers – equivalent to filling over seven 25m swimming pools*

LLMs Already Have A Spam Problem

This is pretty well documented. LLMs seem to be driven at least in part by the sheer volume of mentions in the training data. Everything is ingested and taken as read.

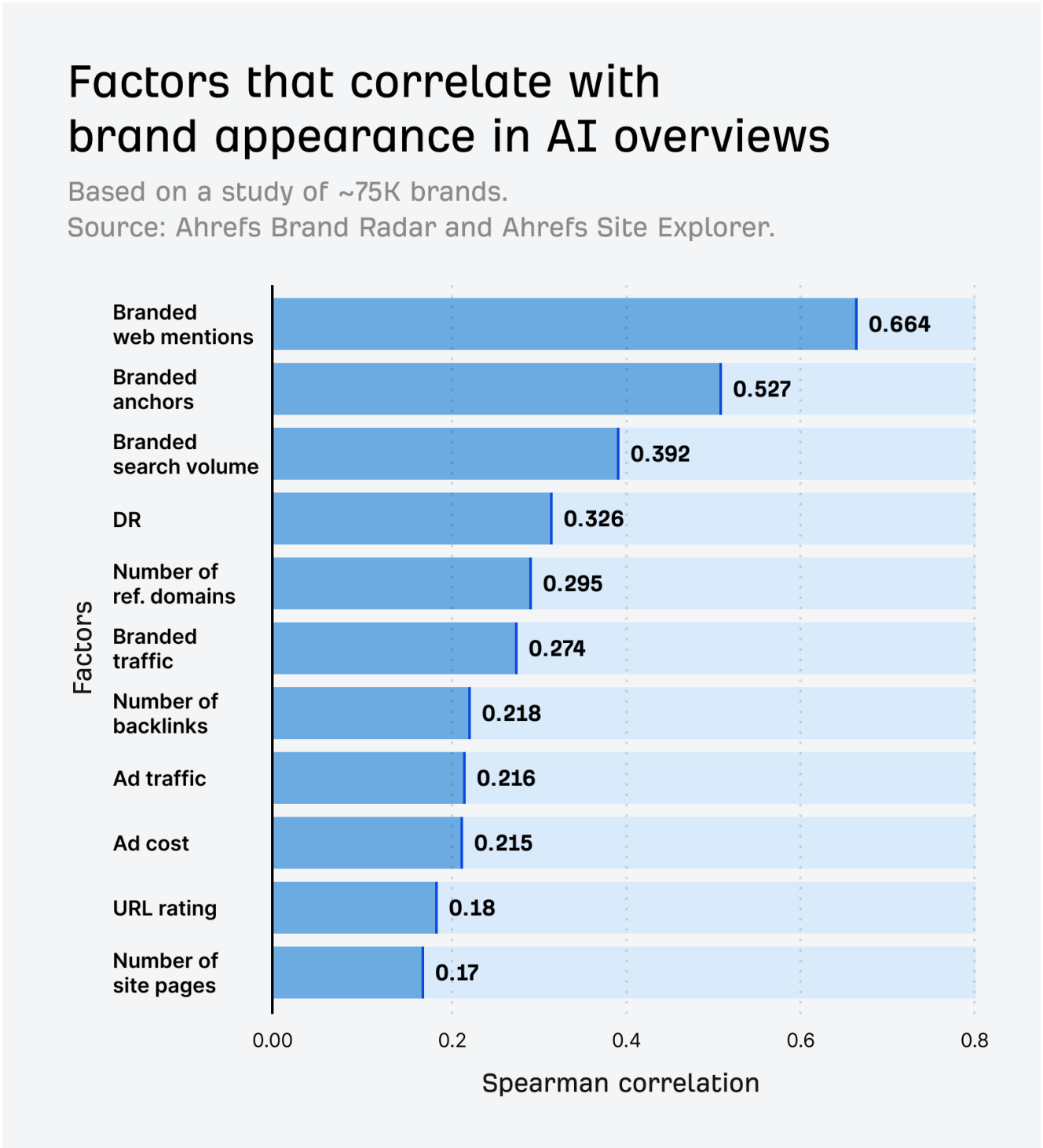


Image Credit: Harry Clarkson-Bennett

When you add a line in your footer describing something you or your business did, it's taken as read. Spammy, low-quality tactics work more effectively than heavy lifting.

Ideally, we wouldn't live in a world where low-lift shit outperforms proper marketing efforts. But here we are.

Like in 2012, "best" lists are on the tip of everyone's tongue. Basic SEO is making a comeback because that's what is

currently working in LLMs. Paid placements, reciprocal link exchanges. You name it.

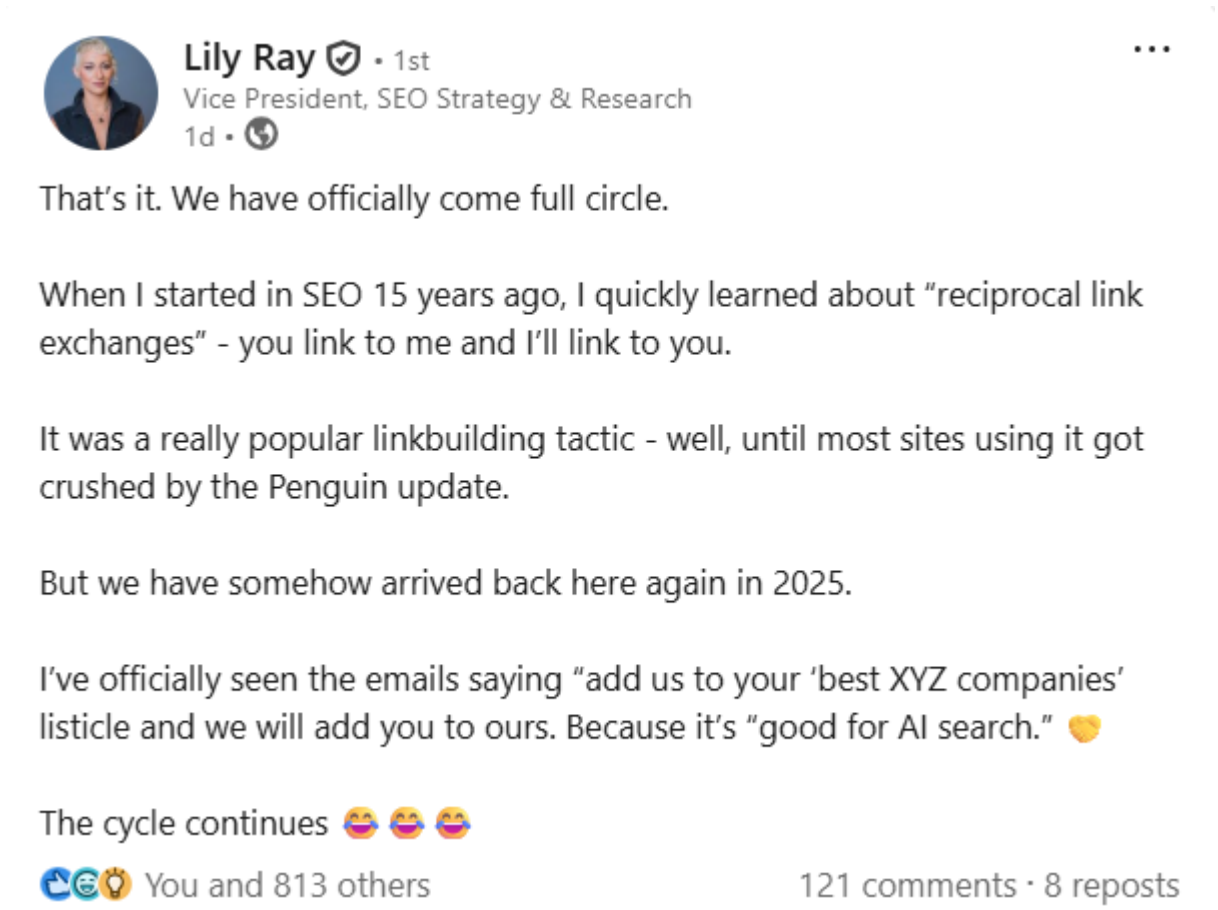


Image Credit: Harry Clarkson-Bennett

If it's half-arsed, it's making a comeback.

As these models rely on Google's index for searches that the model cannot confidently answer (RAG), Google's spam engine matters more than ever. In the same way that I think publishers need to take a stand against big tech and AI, Google needs to step up and take this seriously.

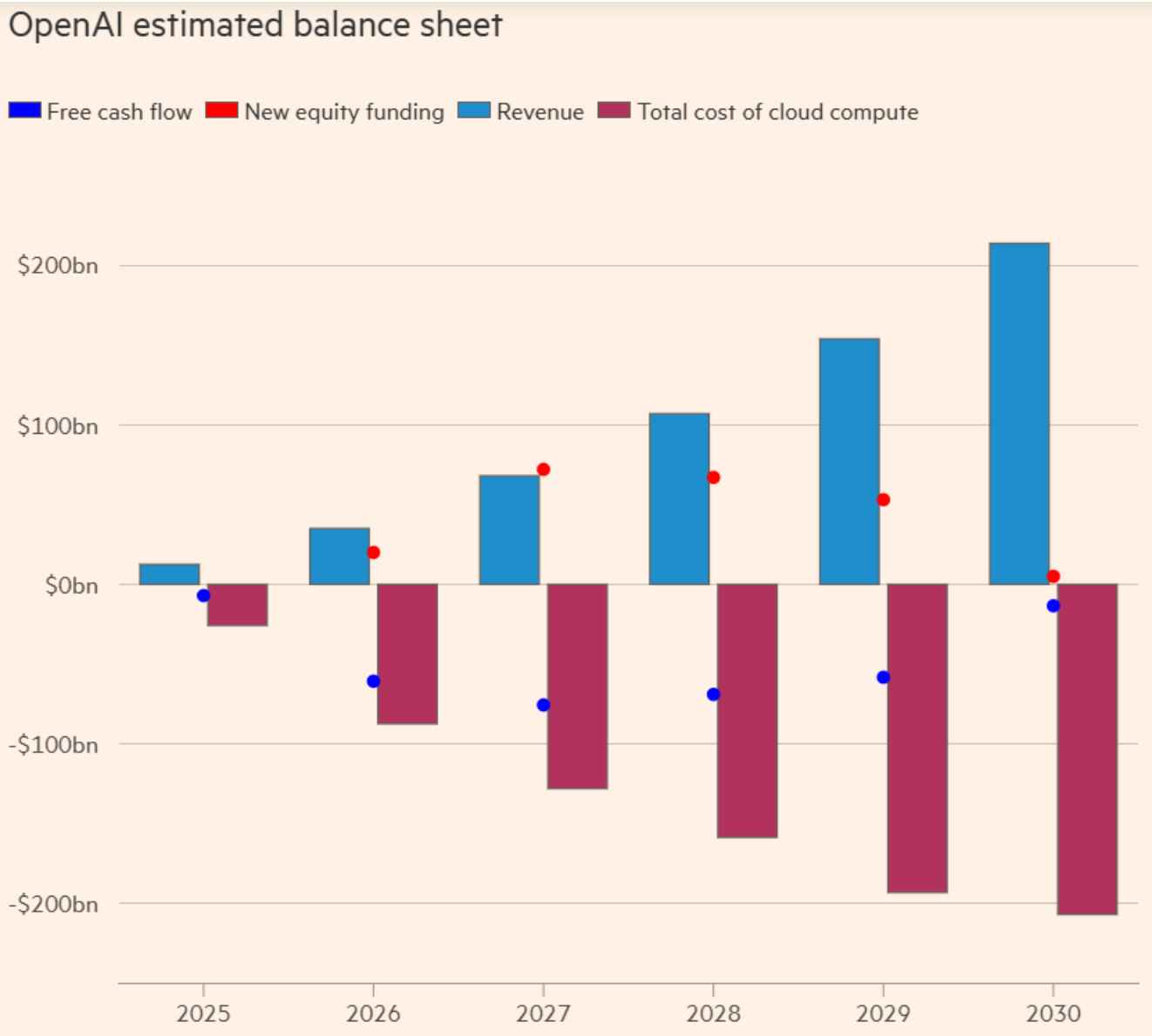
I'm Not Sure Anyone Is Going To…

I'm not even sure they want to right now. OpenAI has signed some pretty extraordinary contracts, and its revenue is light-years away from where it needs to be. And Google's CAPEX expenditure is through the roof.

So, things like quality and accuracy are not at the top of the list. Consumer and investor confidence is not that high. They

need to make some money. And private companies can be a bit laissez-faire when it comes to reporting on revenue and profits.

According to HSBC, OpenAI needs to raise at least $207 billion by 2030 so it can **continue to lose money**. Being described as ‘a money pit with a website on top’ isn’t a great look.



New funding has to be thrown at data centres (Image Credit: Harry Clarkson-Bennett)

Let’s see them post-hoc rationalize their way out of this one. That’s it. Thank you for reading and subscribing to my last update of the year. Certainly been a year.

More Resources:

This post was originally published on Leadership in SEO.

Featured Image: Khaohom Mali/Shutterstock

https://www.searchenginejournal.com/why-googles-spam-problem-is-getting-worse/563258/