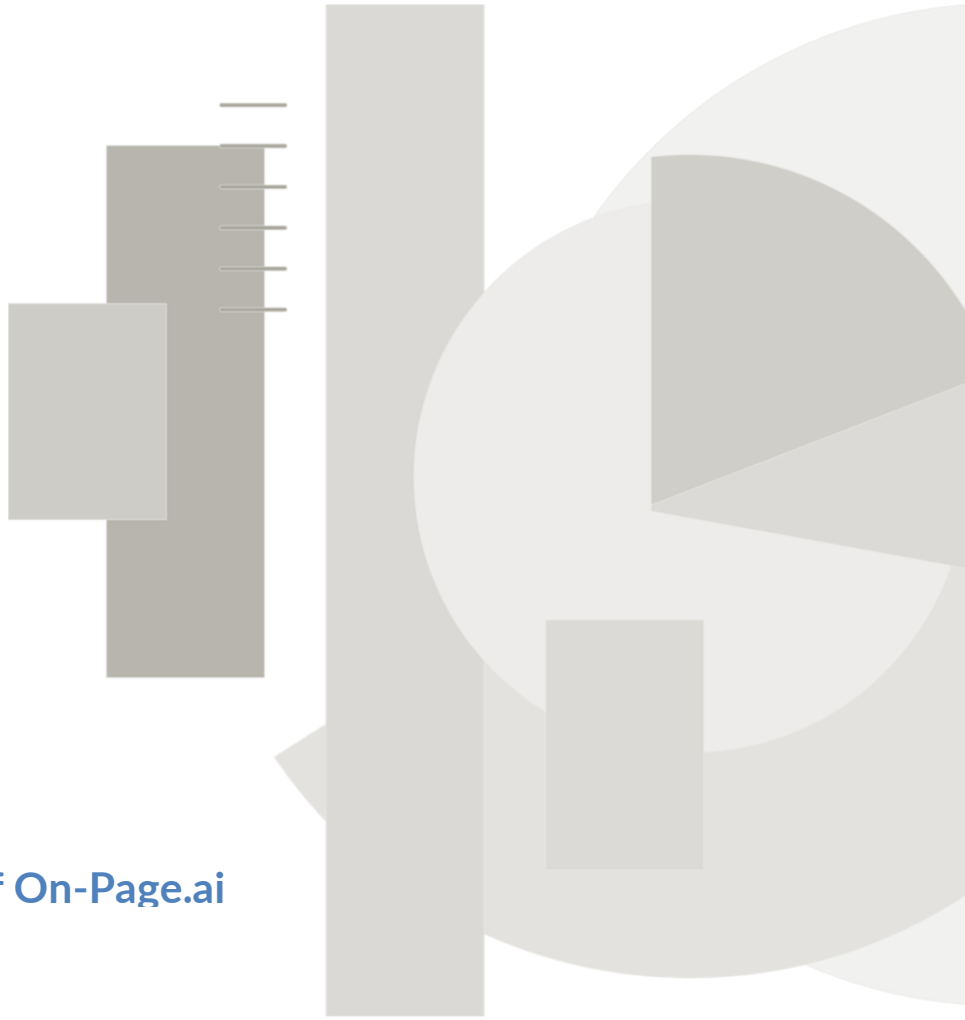
On-Page

Google's Helpful Content Update

Full Review, Analysis and Recovery



By Eric Lancheres

SEO Researcher and Founder of On-Page.ai

1. Google Helpful Content Update Timeline

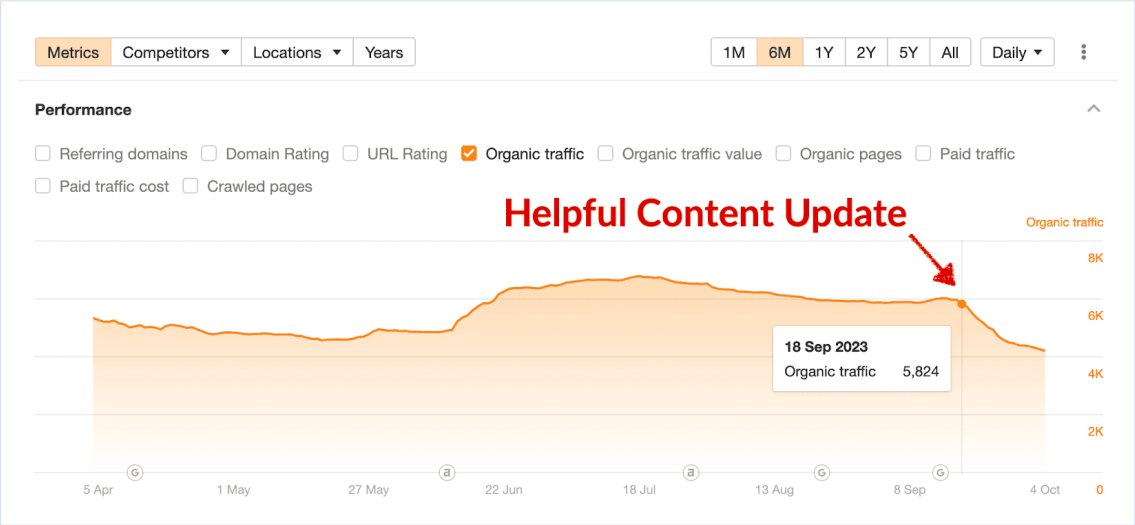
From Rollout To Impact

In the ever-changing landscape of SEO, some updates pass by almost unnoticed while others leave an indelible mark. The Helpful Content Update, also known as HCU, stands out as one of the most impactful updates in recent times.

Launched officially on September 14th, 2023, the most significant impacts from the Helpful Content Update were felt 4 days later on September 18th. The rollout concluded on September 28th, 2023. (Source)

It's important not to confuse this with the August Core Update, which ended on September 7th, 2023 or the October Core update which began Oct 5th, 2023. (Source)

So, if you're among those who experienced a sudden and significant drop in traffic around September 18th (See below for an example), rest assured—you've come to the right place to get answers and solutions.



Ahref organic traffic graph of a website affected by the Helpful Content Update. A sharp traffic drop appears on Sep 18th and continues downward.

2. Official Statements About The Google Helpful Content Update

(Questions & Hints)

While I agree with many search engine experts that Google will sometimes present slightly misleading information to the SEO community, I still believe there is some value in reviewing the official documentations for clues. Here are some official statements about the Helpful Content Update:

1. "*The system generates a site-wide signal*" (Source)

A site-wide signal is a ranking signal that impacts the entire website's performance.

The ranking score for each individual page is derived from a combination of its own score and that of the website hosting it. As a result, **even an excellent individual page may underperform if the overall HCU site-wide score is low.**

2. "The helpful content system aims to better reward content where visitors feel they've had a satisfying experience, while content that doesn't meet a visitor's expectations won't perform as well." (Source)

They are saying that they are focusing on "user accomplishment" which has historically been measured through diverse method ranging from tracking search result clicks to using web data from various third-party sources (Chrome, Android, etc).

3. "This classifier process is entirely automated, using a machine-learning model." (Source)

This is a VERY important new element. This says they are using an **AI classifier to process the data** for the Helpful Content Update.

4. "Our classifier runs continuously, allowing it to monitor newly-launched sites and existing ones." (Source)

The AI classifier is constantly running which means that **you can recover at anytime** and there is no need to wait for another Google Helpful Update in order to recover. As soon as Google recalculates your score, recovery should follow.

5. "Sites identified by this system may find the signal applied to them over a period of months" (Source)

This suggests that they might need to accumulate enough data (user signals) to accurately calculate a score. Alternatively, it may also mean they need time to crawl an entire website in order to properly calculate the signal.

Google provided a long list of questions to ask yourself when creating content to self-assess if it is helpful. While these front-facing human can serve as useful guidelines, they do not necessarily represent how the algorithm functions.

Some of the questions include:

6. "Does the content present information in a way that makes you want to trust it, such as clear sourcing, evidence of the expertise involved, background about the author or the site that publishes it, such as through links to an author page or a site's About page?"

7. "Is this content written or reviewed by an expert or enthusiast who demonstrably knows the topic well?"

8 . "Does your content clearly demonstrate first-hand expertise and a depth of knowledge (for example, expertise that comes from having actually used a product or service, or visiting a place)?"

9. "Do you have an existing or intended audience for your business or site that would find the content useful if they came directly to you?"

10. "Does your content leave readers feeling like they need to search again to get better information from other sources?"

11. "Do bylines lead to further information about the author or authors involved, giving background about them and the areas they write about?"

And they mention this important bit:

12. "The "why" should be that you're creating content primarily to help people, content that is useful to visitors if they come to your site directly. If you're doing this, you're aligning with E-E-A-T generally and what our core ranking systems seek to reward." near the end. (Source)

The 'why' is interesting because it hints that the at the idea of direct visitors for a website. A subject which we will expand on further down in the report.

3. Calculating The Helpful Content Score

(Using a new AI classifier)

The Helpful Content Score is determined by a process that uses AI to analyze the helpfulness of a page. (Source)

First introduced back in August 2022, the initial version of the Helpful Content Update was relatively mild and served only to detect the most egregious offenders, creating mass-low-quality AI content.

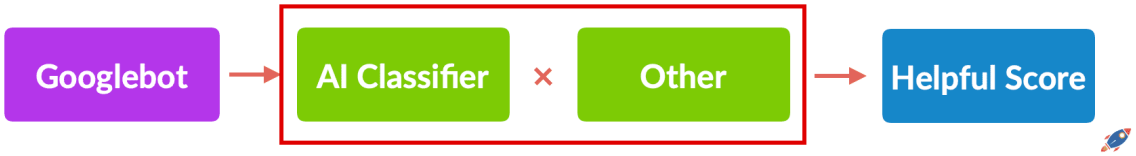
Fast forward a year, and the new Helpful AI classifier is drastically different and a huge improvement over the original. Unlike earlier technologies that relied on predetermined quality indicators, **the AI classifier now understands the content.**

For example, the latest search engine results seem to indicate that the new AI classifier can locate *multiple* answers provided on a single page.

We also have tentative evidence that the AI classifier might be smart enough to understand that an author specializing in business may lack the qualifications to provide medical advice, even if that author is considered trusted in business-related domains.

While Google keeps the process behind the classifier a secret, we can make a few assumptions on it's functionality.

Assumption #1: The Helpful Content AI Classifier Has Help



The classifier is likely **accompanied by other modifiers and ranking factors in order to arrive at a final helpful score.** It would be unusual for Google to just take the results straight from an AI classifier, without any other filters or modification to the result.

There are likely other signals (ranking factors) at play when categorizing helpful content from around the web.

Assumption #2: Helpful Content AI Classifier Is Text Based

The AI classifier processes only text.

Large language models are now commonplace and have been shown ready for production. While Google has shown development of a multi-modal AI model, **deploying such a model to analyze images, layout, and text simultaneously would be resource-prohibitive on a worldwide scale.**

While I'm sure Google *does* care about images (more on this later), the decision to categorize a page as helpful is likely influenced more by text rather than imagery.

Assumption #3: Helpful Content AI Classifier Limitations

With years of experience with AI models under my belt, I began to imagine how Google engineers might go about developing an AI classifier designed to evaluate the helpfulness of a webpage.

This is when I had my first breakthrough:

"The AI classifier cannot process the entire page."

If Google is using an AI classifier (and their official documentation says they are), then they are heavily bound by memory and processing size. In comparison, the world's most popular AI model, OpenAI's GPT 3.5-Turbo, has a limitation of 4096 tokens.

Another example is Llama 2 from Meta, which also has a 2048 token limit. In fact, it seems like MOST modern AI models currently have a limited token limit. Larger models do exist however they consume quite a lot more memory and it would be impractical for Google to process the entire web with a larger model.

Google is likely using a derivative or fine-tuned version of the Palm 2 to evaluate pages.

The Palm 2 model has an input token limit 8196 which is equivalent to approximately 6147 words.

When using AI language models, you need to provide it with detailed instructions that can easily take up 300 tokens, so in the best case scenario, we'd see approximately 7896 tokens allocated to the input of the text. This would equate to approximately a maximum input of 5922 words.

This is the best case scenario *if* Google was using their most compute-intensive model to crawl their entire web.

This is highly unlikely.

Much more likely is that they are using a smaller AI model, similar in size to their Palm 2 Gecko model, which has a 3072 token input limit (*they aren't using this exact model as it is designed for embeddings however it does give us an idea of the different variations of the models they are producing internally*).

This model is MUCH faster, consumes a fraction of the memory and COULD feasibly process the entire web's content. **An educated guess would be that the AI model Google developed for the Helpful Content Update uses a token input limit from 2048 to 4096, with 3072 being a very possible middle ground token limit.**

In practical terms, this equates to approximately a raw input of 2304 words and if we imagined a very compact instruction set (~300 words) alongside, we could realistically see a maximum of 2000 words as the maximum input when crawling the web.

I believe that the Helpful Content AI Classifier only looks at the first 2000 words of a page in order to asses if the page is helpful.

While this is supported by my analysis, please note that this is just a speculative guess.

However, if we function under the assumption that the AI classifier has limitations, then it means that if you are creating lengthy articles and all your "trust signals" are BELOW the content, they could be missed entirely by the Helpful Content Update AI.

For example, references and an author bio below a 6000 word article could be missed.

I believe it is important to demonstrate helpfulness at the top of the article and if possible, keep the length of the content on the shorter side WHILE preserving all the important information. Essentially, writing with less fluff.

Assumption #4: Helpful Content AI Classifier Works On Articles

While Google's algorithm is used to rank to all pages on the internet, it is implied that **Google classifies content & websites by type.**

(This is under the assumption that Google has ALREADY classified the internet. It is possible that the new AI Helpful classifier is also being used to re-classify the web. I'll explain this further in Theory #3)

For example, the "Product Review Update" targets... reviews.

Here are some general classifications commonly referred to on the web:

- Homepage
- Category page / Navigational Page
- Product Page
- Article (Informative, Review, Opinion, News, Informative)
- Resource Page
- Video Page
- Discussion Pages
- Social Pages
- Other

Google likely uses Schema-based types to determine the content type.

- Article
- Event
- Local Business
- About Page
- Profile Page
- Product Page
- FAQ Page

... and many more.

While we don't know the exact categorization that Google uses internally, the point is that it just wouldn't make sense to apply the Helpful Content AI classifier to every page.

It wouldn't make sense to apply it to a video page or a homepage. Nor would it make sense to try to evaluate how helpful a category page is!

Those pages are important yet should not be evaluated by the same Helpful Content AI classifier.

In fact, the only reasonable way to get meaningful results would be to use the Helpful Content AI classifier on pages it was trained to assess.

I believe the AI classifier is used for **article type pages** and *maybe* for **forum/discussion pages.**

In contrast, I do not believe it is currently being used to evaluate your homepage, your category pages, video pages, ecommerce products and even... resources pages.

One anomaly I noticed is that many resource pages remained unaffected by the Helpful Content Update. These were often found on custom CMS / custom coded websites and provided value without any trust signals.

As such, **I believe the Helpful Content AI classifier was built and trained with unhelpful WordPress articles as their main target.** This is not to say that other pages were unaffected but instead, I believe that low quality WordPress sites were the main target.

(A reminder that this the Helpful Content Update is a site wide signal which means that if the bulk of your pages are deemed unhelpful, then it could potentially drop your homepage rankings, even if your homepage is not specifically evaluated by the Helpful Content AI classifier)

4. How The Google Helpful Content Update Alters The Google Algorithm

(The Severity Of The Impact)

The Google helpful content update adds a new "Helpful Score" metric to the Google algorithm that determines how helpful your website is to users.

Let me explain.

If we take an **overly simplified model** of the Google algorithm,

(The Google ranking algorithm is comprised of hundreds of ranking factors that are carefully weighed against each other to dictate the search engine results. While we describe ranking factors with common representative names such as links, domain authority, content relevance, speed, user experience, Google uses drastically different terminology within their algorithm. For the sake of understanding what is happening with the Helpful Content update, it is not necessary to use the same terminology as Google)

Grouping all the traditional ranking factors inside together, we can imagine:

EXAMPLE:
Website #1: [Google Ranking Algorithm] = 94
Website #2: [Google Ranking Algorithm] = 87
Website #3: [Google Ranking Algorithm] = 67
Website #4: [Google Ranking Algorithm] = 59
Website #5: [Google Ranking Algorithm] = 58

At a basic level, the higher your ranking score, the higher you rank in the search results.

The Helpful Content update introduces a new variable, a "Helpful Score" for your entire website that measures the helpfulness of your website. It might look like this:

[Google Ranking Algorithm] x [Helpful Score] = Ranking Score

This is a significant ranking factor that impacts your entire website which is why some webmasters have been experiencing severe site-wide traffic drops.

Using our previous example, if you own website #1 and your helpfulness score is determined to be 70% (0.7), then you could be facing a 30% drop in ranking score.

EXAMPLE:
Website #2: [Google Ranking Algorithm] = 87
Website #3: [Google Ranking Algorithm] = 67
Website #1: [Google Ranking Algorithm] = 94 x 0.7 = 65.8
Website #4: [Google Ranking Algorithm] = 59
Website #5: [Google Ranking Algorithm] = 58

Assuming all the competitors scored a perfect 100% score as Helpful Content Score and you received 70%, then it would be enough to make your rankings drop to position #3.

Now let's assume your competitors received Helpful Scores ranging from 100% to 30% (1.0 to 0.3), we can see some significant changes in the search results.

EXAMPLE:
Website #3: [Google Ranking Algorithm] = 67 x 1.0 = 67.0
Website #1: [Google Ranking Algorithm] = 94 x 0.7 = 65.8
Website #4: [Google Ranking Algorithm] = 59 x 0.5 = 29.5
Website #2: [Google Ranking Algorithm] = 87 x 0.3 = 26.1
Website #5: [Google Ranking Algorithm] = 58 x 0.3 = 17.4

As you can see, the Helpful Content Score plays a critical role in determining your final score and your ranking position in the search engine results.

Note: I suspect that the helpful score can only reach a maximum of "1" and therefore cannot reward sites, only penalize them. More on this later.

5. Statistics, Charts, Figures

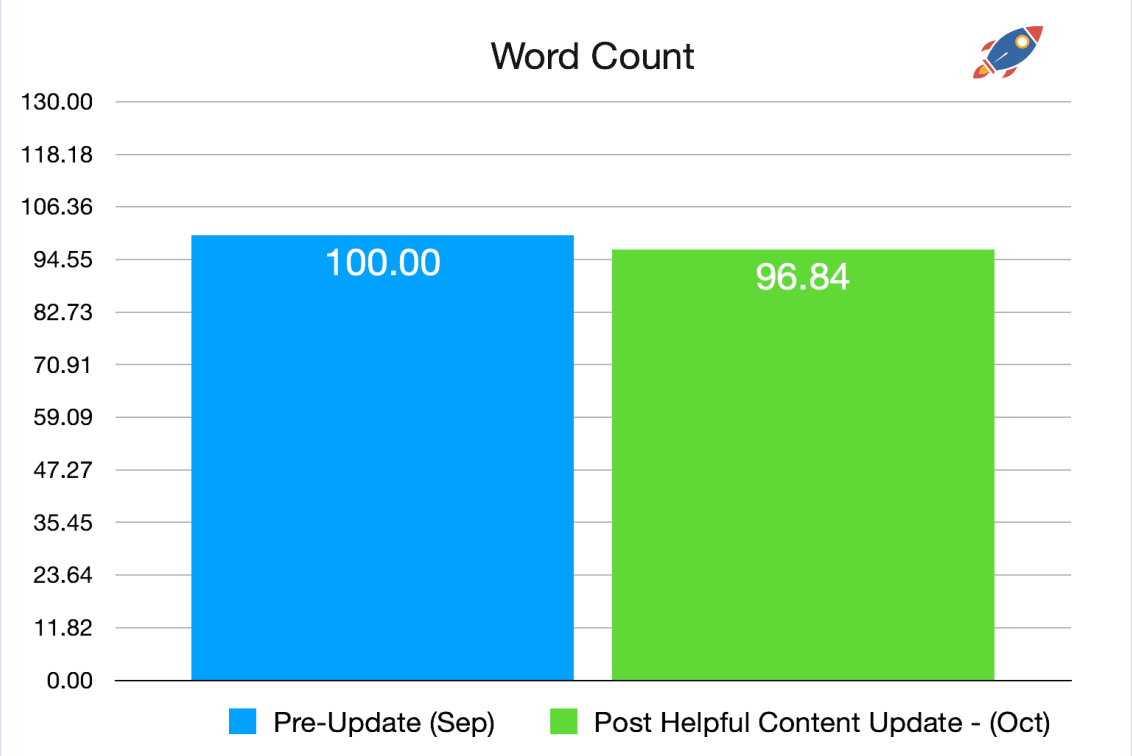
(Before & After Data)

Fortunately, I have access to quite a lot of data and as we have been tracking search engine fluctuations for years for previous Google Core Updates.

Here is the **result of over 100,000 data points comparing the BEFORE and AFTER during the Helpful Content Update**. Please note that correlation does not mean causation. These charts provide a snapshot into the ever-changing search landscape.

Before: Early September 2023
After: Early October 2023

Average Page #1 Word Count



Historically, the average word count has been fairly stable. This is why I believe that a **3.26% reduction in average word count** after the September 2023 Helpful Content Update is a little unusual.

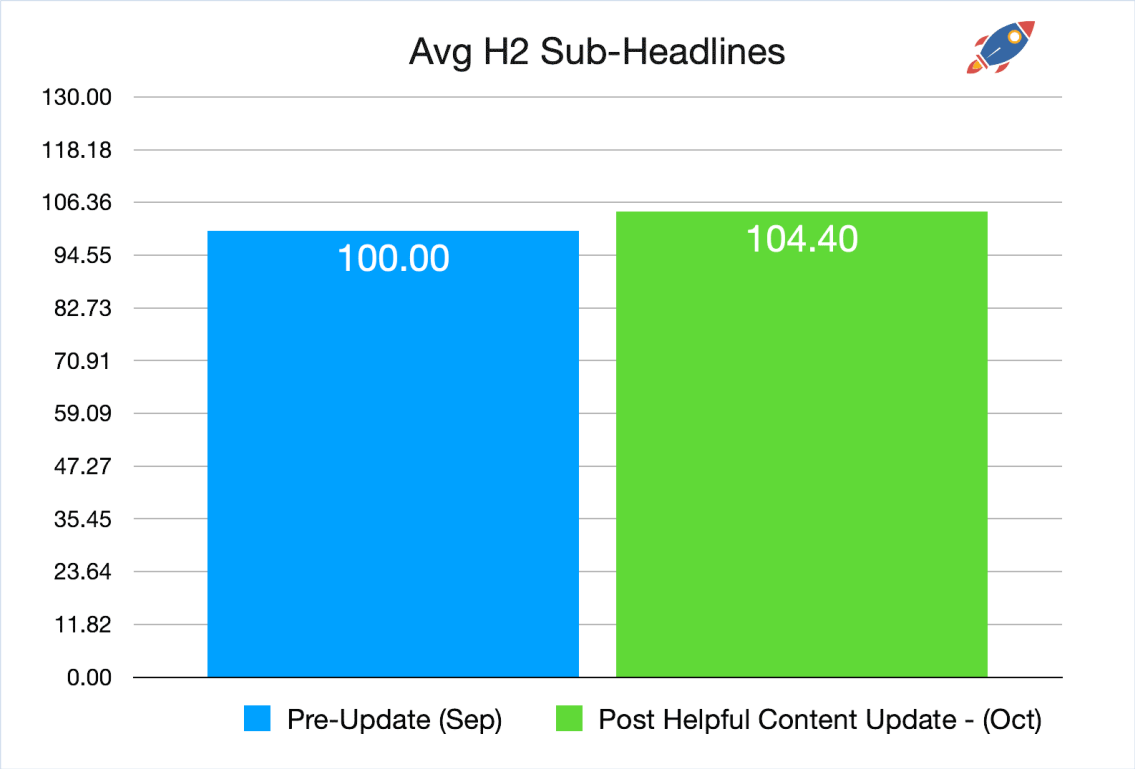
This indicates that, on average, there is more short form content ranking in the search engine results.

Anecdotally, it is the first time in a very long time (possibly ever) that I have seen top ranking results in the 300-400 word range. For context, before the Helpful Content Update, the average word count hovered around the ~1400 word mark and results with less than 800 words were exceedingly rare.

To see a 309 word article from 2014 rank in the #1 position for a keyword is... unusual to say the least!

Note: For *some* keywords, word count actually increased however on average, we observed it to be lower, likely due to unusually short content in some instances which brought down the average.

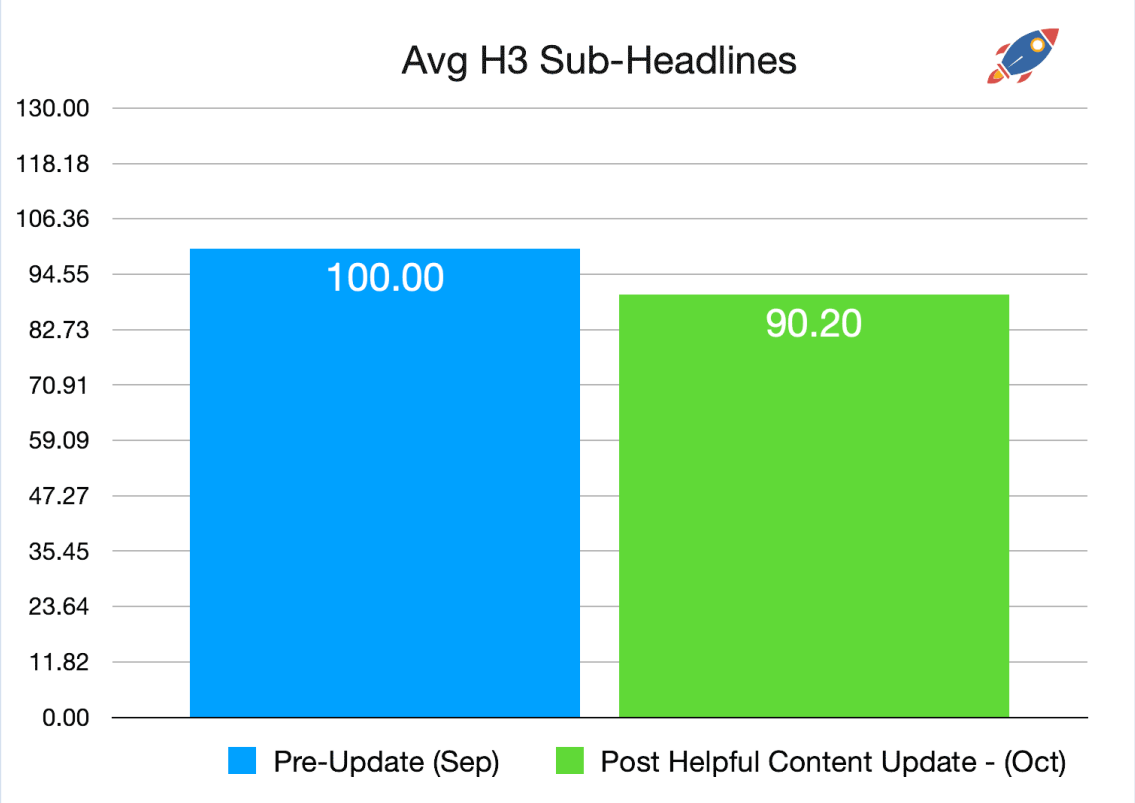
Average Quantity Of H2 Sub-Headlines



The average quantity of H2 sub-headlines on page 1 of Google results increased by 4%. This is not as meaningful as the overall word count which traditionally has been very stable.

It is, however, **odd to see an increase of H2 sub-headlines while the overall word count is decreasing.** (You would assume that it would decrease in a linear fashion alongside the word count.)

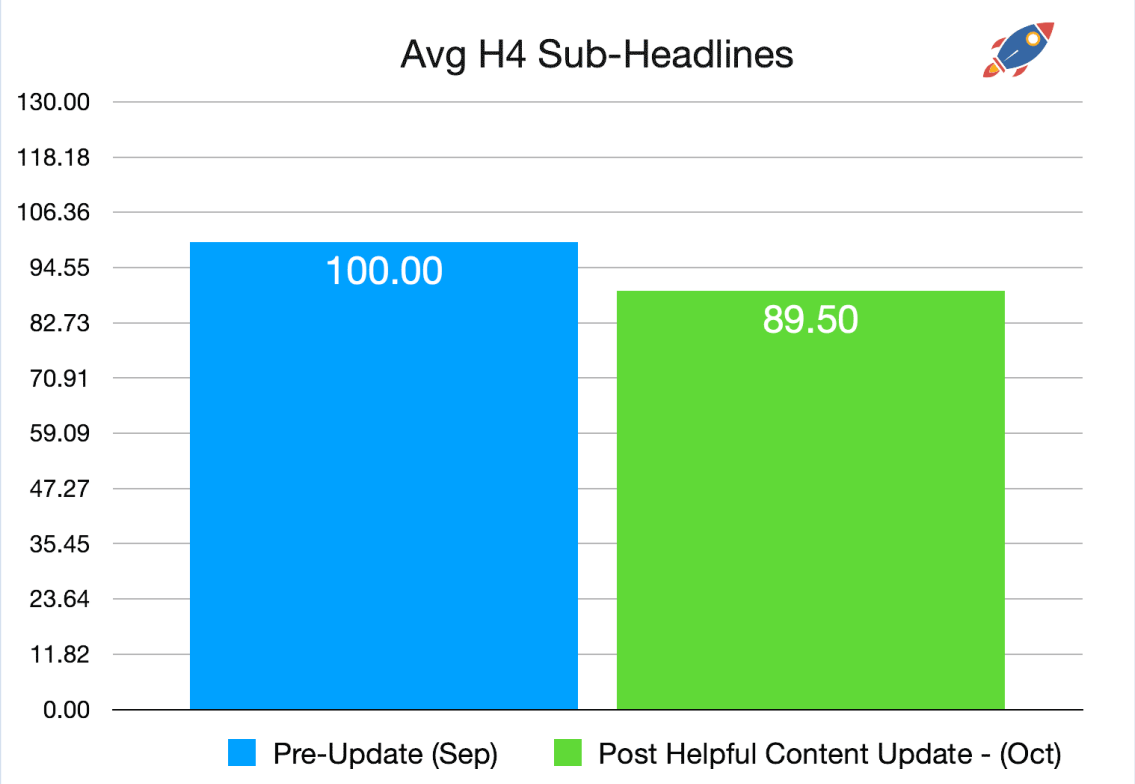
Average Quantity Of H3 Sub-Headlines



The average quantity of **H3 sub-headlines on page 1 of Google results decreased by nearly 10%.** This is a substantial drop which might be attributed to some results being quite short.

It appears that larger "cover everything" articles are being replaced with narrowly focused articles that don't go into depth into sub-topics.

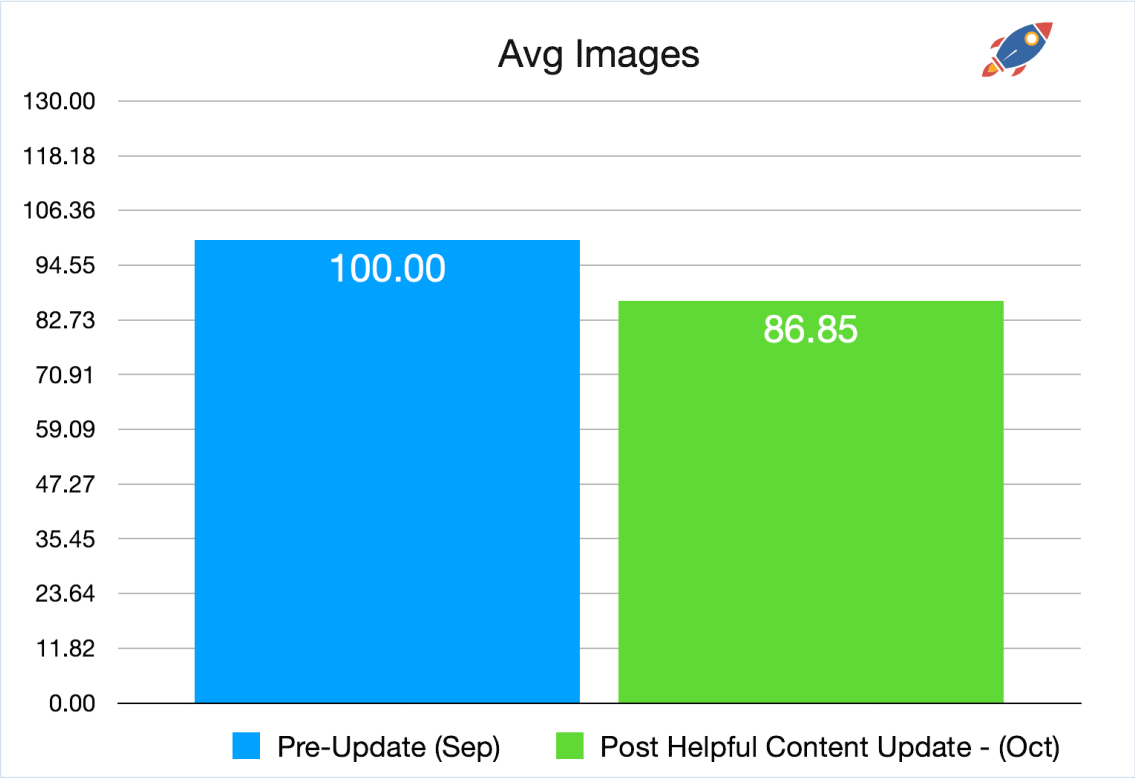
Average Quantity Of H4 Sub-Headlines



The trend accentuates itself with the **average quantity of H4 sub-headlines diminishing by 10.5%.** While H4 sub-headlines are rarely used in comparison to H2 and H3 sub-headlines, they indicate, once again, that content that dives deep into sub-topics is not as present within the search results.

It seems as, at least from the data, narrowly focused H2 content seems to be winning.

Average Quantity Of Images Within Main Content

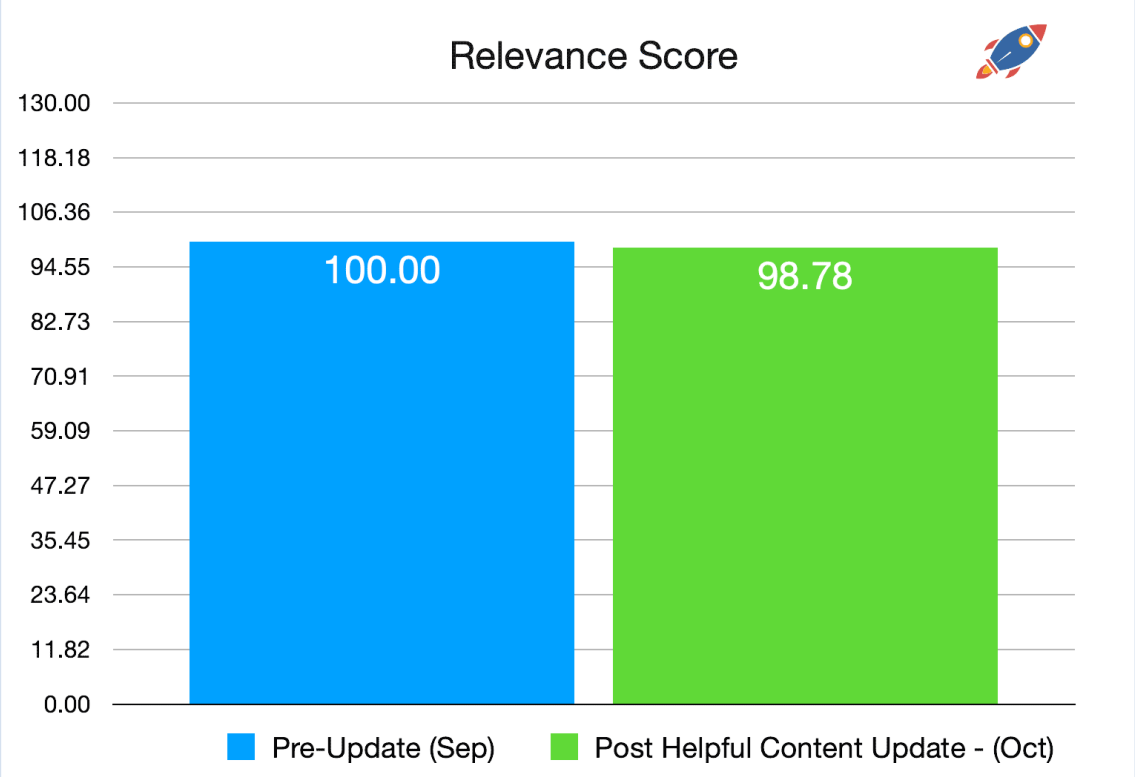


This was a HUGE surprise as I originally expected the average quantity of images within the main content to increase. Reading the helpful content guidelines would suggest Google would have been rewarding pages with large quantities of images (hopefully original images to prove they have first hand experience) yet the opposite is true.

Images within the main content dropped by 13.15% on average.

Perhaps this is due to more shorter form content ranking within the search engine results or maybe it's just that longer, in depth articles aren't being favored as much.

Relevance Score

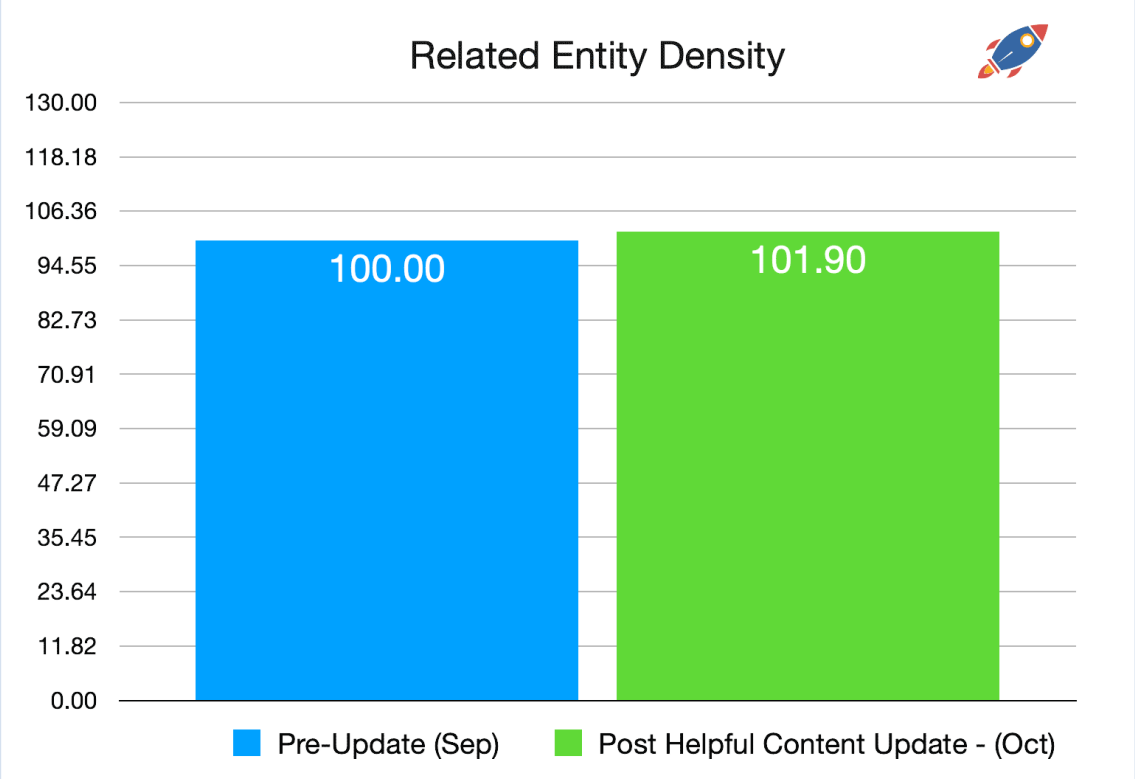


The relevance score is a unique On-Page.ai metric that mimics modern search engines to measures the relevance of the document for a specific term based on the search term, related entities, word count and entity density.

It aims to help you optimize your page by adding related entities to your content while minimizing fluff and irrelevant text.

The slight 1.2% in reduce of relevance score indicates that Google is favoring *slightly* less SEO optimized documents after the update.

Entity Density



Related entity density measures how often related entities are found within the text in relation to filler / other words. **While the overall raw entity count decreased due to the word count diminishing, the entity density actually increased by 1.9%.**

This indicates that Google's algorithm still reads, understands and seeks out content with related entities, regardless of the length.

6. In-Depth Site Analysis

Trends & Observations

The process of reviewing hundreds of sites manually is not very scientific yet I feel it remains one of the best ways to identify trends, develop new ideas and understand a new algorithm update. It is often by identifying the exceptions that we learn the most and can validate our general theories on the current algorithm.

Although over 117 sites were manually reviewed for trends over a period of 1 week, I didn't think it would be useful to include a breakdown of every single site. This is just a tiny fraction of exemplary sites that stood out.

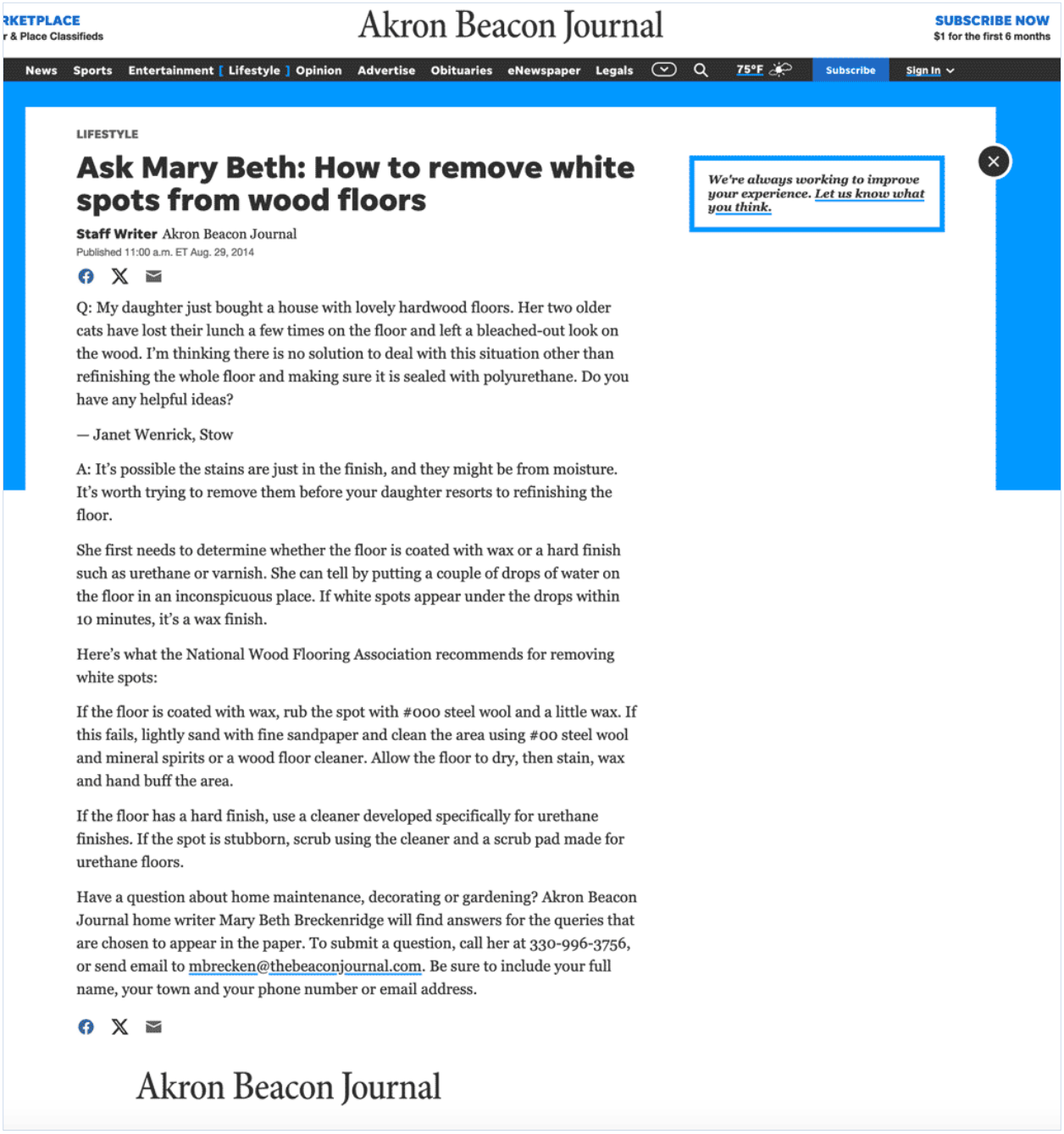
Here's what I discovered:

Super Short & Outdated Content

This is the Ahref organic traffic graph of a **309 word article from 2014 that saw a significant increase** after the Helpful Content Update for the term: "*how to remove white spots from wood floors*".

Wow.

Here's how it looks:



The article isn't bad per say but what is interesting is that it isn't your traditional SEO optimized article.

- There are no sub-headlines.
- There are no images.
- It is a very short article (309 words).
- The article hasn't been updated in 9 years.
- No subject / topic coverage beyond the exact question.
- No fancy formatting.
- The author By Line says: "Staff Writer".
- It's on a general site that does not specialize in hardwood floors.
- There are no contextual links within the content.

And yet, it ranks!

I believe this demonstrates the capabilities of the new AI classifier that can now read and *understand* content, seeking out "helpfulness" trust signals. It also points towards other ranking signals being at play beyond just the AI classifier.

So what can we observe?

1. Real Names

We have"Janet Wenrick" mentioned within the article and the person answering, "Mary Beth Breckenridge". While the older algorithm would have likely only looked at the ByLine, seen "Staff Writer" and then left... I suspect the new Helpful Content Classifier can process the entire content to understand if the person writing is reputable. It is likely that the Helpful Content algorithm sees identifiable names and usernames as a trust signal.

2. Phone number AND Email

It is rare to list personal information within an article so this one stands out as the exception. It is **very likely that having a phone number and email will increase the chances of the article as Helpful by the AI classifier.**

3. Integrated Quote / Reference

Previously, the "*Here's what the National Wood Flooring Association*" part would have been completely missed by the old algorithm. Yet, this is a valid reference that provides trust and authority to the answer. This is, yet again, something only an AI could pick up and **having a reference from the National Wood Flooring Association likely plays a role in determining how trustworthy and helpful the answer is.**

4. Short article length might be a strength

Remember how I discussed that the AI classifier is very likely limited by how much content it can process?

Having a very short, **309 word article means that the AI classifier can easily crawl and understand the entire article**... including the names, information about the author, the subtle reference to the wood flooring association.

This information would usually be lost as it is definitely not within the expected spots nor are there any indications within the formatting.

5. Very low fluff

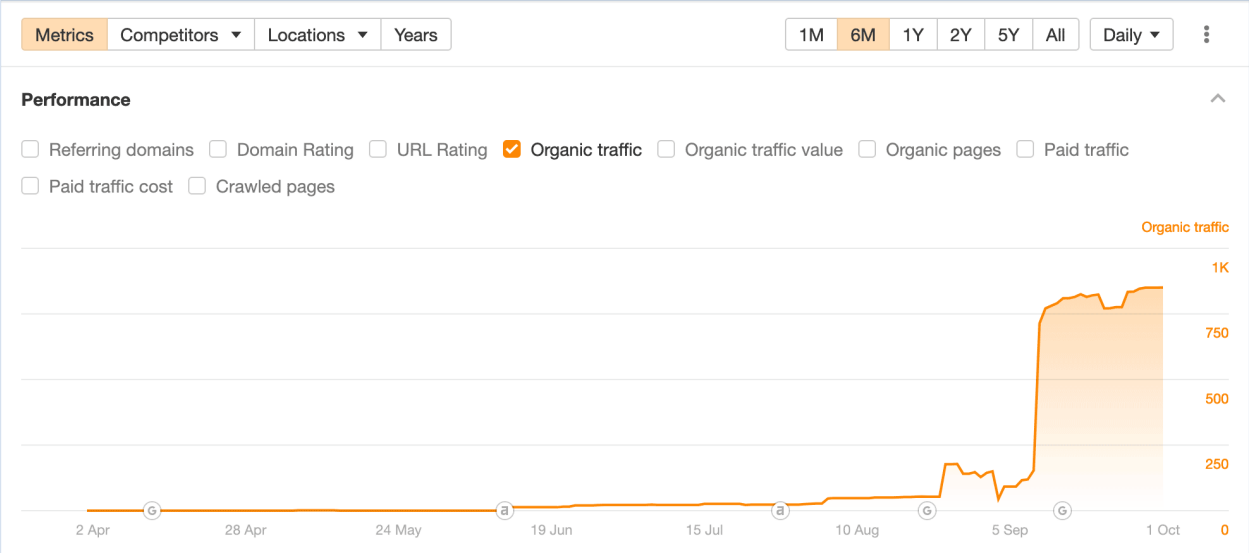
This article focuses exclusively on the topic at hand without deviating. It appears as if **narrowly focused articles might be performing better**, even if they are extremely small.

6. Other trust signals

While I cannot confirm this, after reviewing hundreds of sites, one of the overarching trends is that **it appears as if "returning visitors" might be a trust signal that plays a role in determining a helpfulness score for a website.**

This article resides on a local newspaper site that likely has a very healthy quantity of direct and returning visitors.

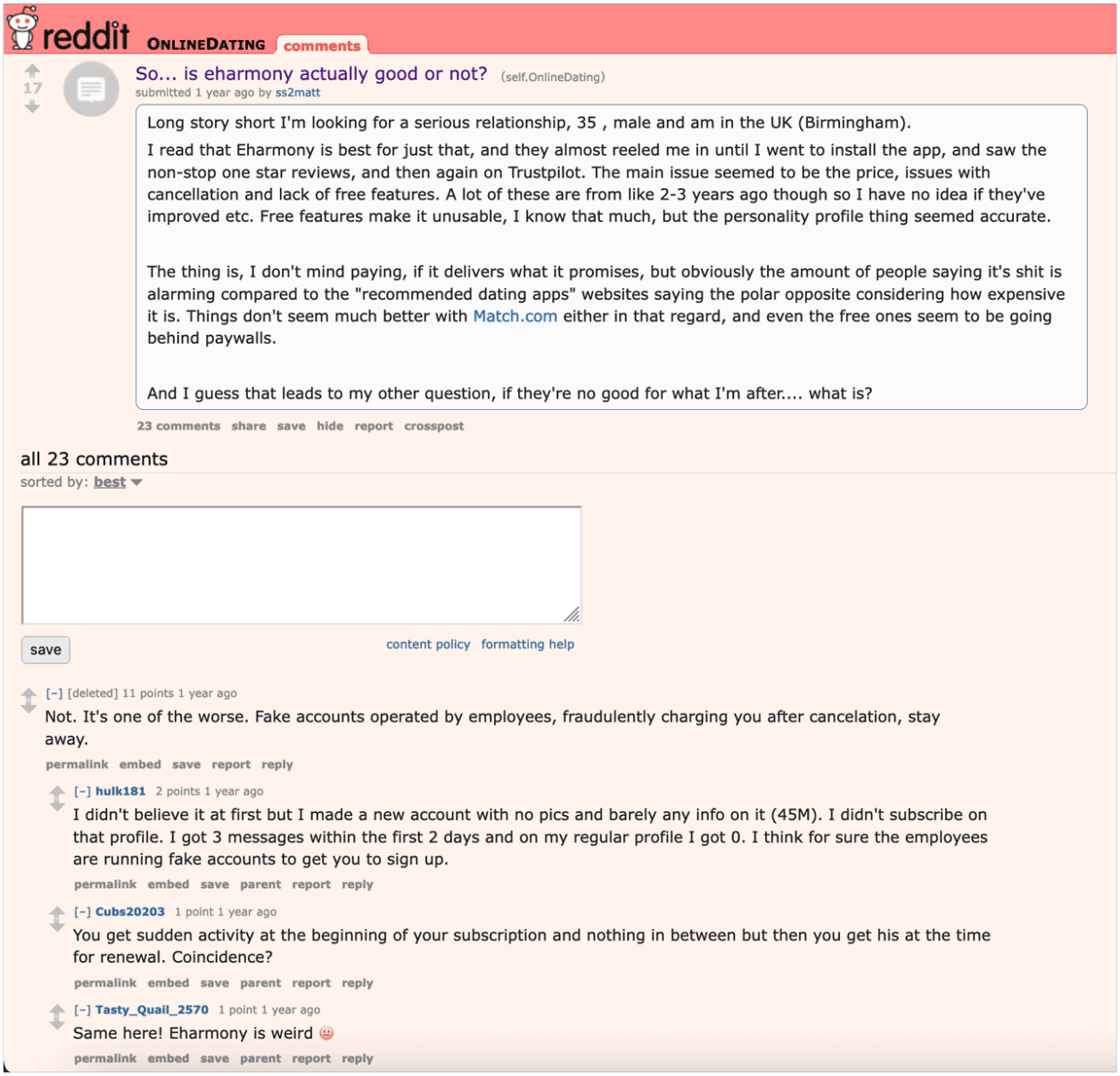
Unusual Reddit Rankings



This is the Ahref organic traffic graph of Reddit article ranking for the term: "*eharmony review*".

While it is widely known within the SEO community that forums, **Reddit and Quora have seen significant visibility increases after the August Core Update** (Not be confused with the Helpful Content Update which occurred a few weeks later). What is interesting is that **the site has maintained the rankings even *after* the Helpful Content Update** meaning that Google must believe these pages are somehow, helpful.

Here's how it looks:



What is VERY interesting about this page is the lack of common SEO optimization for such a competitive keyword.

For example, the title tag, historically one of the most crucial elements for ranking, is:

"*So... is eharmony actually good or not? OnlineDating*"

To my surprise, the seemingly important word "review" doesn't appear at all. In addition, this is more of a sentence rather than a descriptive title.

Even the content itself is not your traditional SEO optimized content. **It's just 432 words of comments about people's experience on eHarmony.**

And yet, it does contain reviews of the eHarmony website and it does rank! Surprisingly, I'd say Google gets this one right.

Of course, I understand it's very easy to dismiss this result because "Reddit ranks for everything" but the question remains...

"*Why THIS page out of millions of Reddit pages?*"

And more importantly,

"*Why DOES Reddit rank for for everything?*"

I believe this exceptionally small and unrelated (by classical SEO standards) result can offer clues into Google's new ranking system.

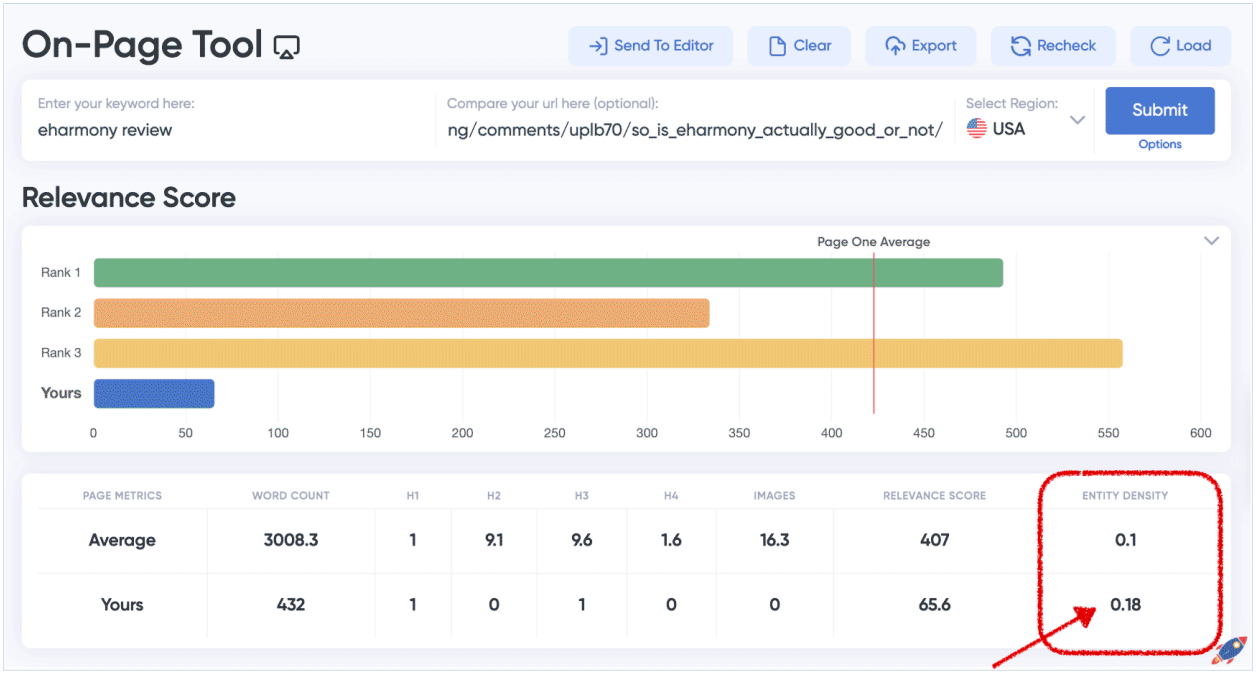
So what can we observe?

1. Names

In the previous example, we highlighted the frequent use of real names however in this example, we have multiple usernames. **It appears as if the Helpful Content Classifier is heavily favoring identifiable names and usernames as a trust signal. This might indicate that Google is favoring different perspectives on a single page.**

This is a recurring theme in many of the helpful pages across the web.

2. High Related Entity Density



While the page itself remains an outlier due to it's minuscule size (and therefore has lower overall page relevance due to the shorter content), **the related entity density is actually significantly HIGHER than the competition.**

At 0.18 entity density compared to 0.1 for the average, it is nearly *double*. This is due to the very narrowly focused topic being discussed within the page.

What's interesting is what ISN'T there:

- There are no mentions of eHarmony's History.
- No instructions on how to sign up or how to use it's features.
- There are no 'deep dives' into sub-topics about eHarmony.

Nothing.

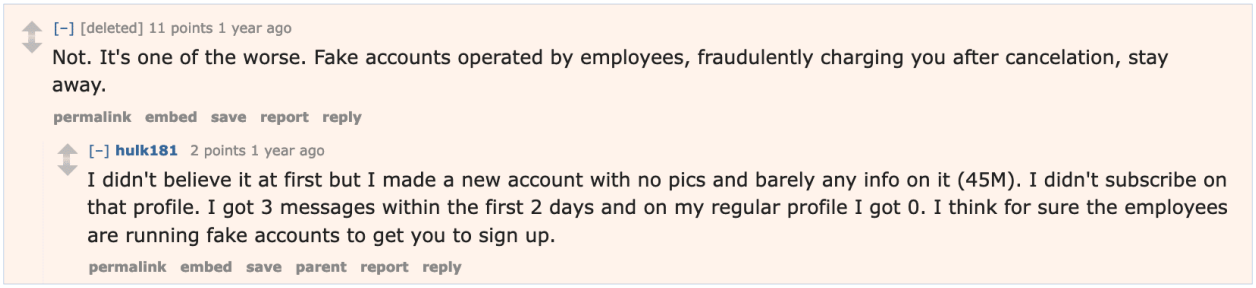
It's just reviews.

And consequently, there is a very high density of related entities such as *eHarmony*, *relationship*, *review*, *price*, *cancellation*, *account*, *app*, *match*, *fake*... and the list goes on.

3. Short content length

As I previously discussed, the AI classifier is very likely limited by how much content it can process and short content allows it to properly assess the entire webpage which it likely deems "helpful".

Previously, older algorithms would struggle to see comments such as:


[–] [deleted] 11 points 1 year ago
Not. It's one of the worse. Fake accounts operated by employees, fraudulently charging you after cancelation, stay away.
permalink embed save report reply
[–] hulk181 2 points 1 year ago
I didn't believe it at first but I made a new account with no pics and barely any info on it (45M). I didn't subscribe on that profile. I got 3 messages within the first 2 days and on my regular profile I got 0. I think for sure the employees are running fake accounts to get you to sign up.
permalink embed save parent report reply

And determine this to be related and useful for the keyword "eHarmony review". There are no mentions of eHarmony and instead, there are only experiences being describe.

Yet, with the new AI understanding the content, it CAN determine that this is related and helpful for the term "eHarmony review".

4. Navigational searches

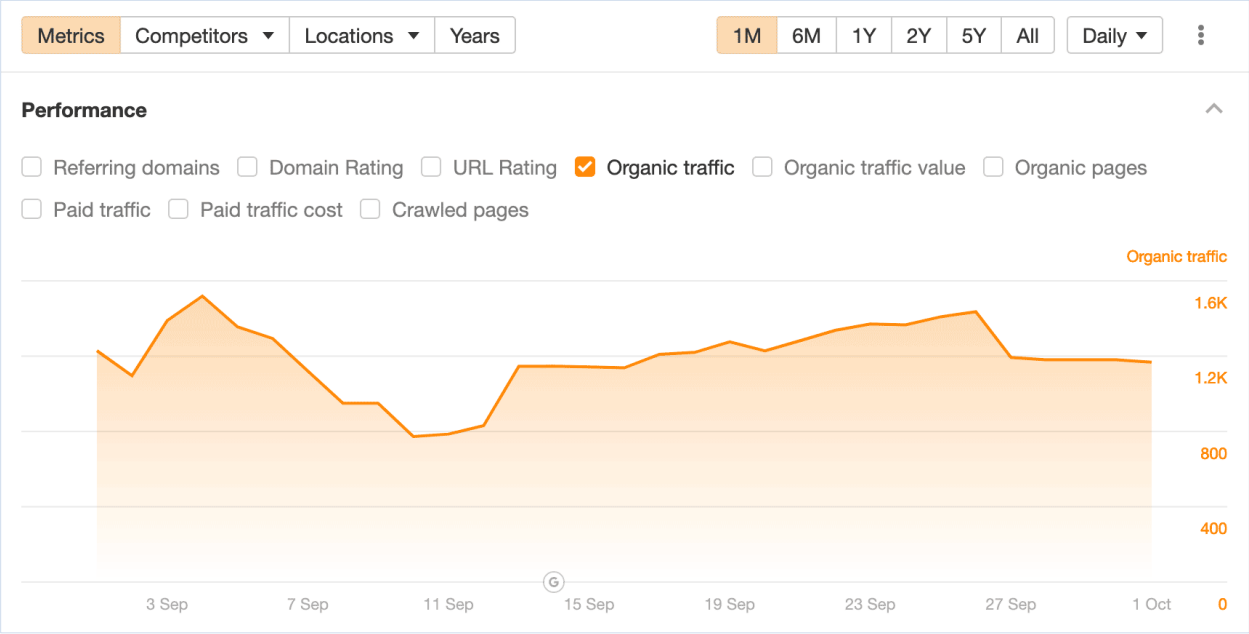
Finally, I believe that one of the "other" signals that work in tandem with the Helpful Content Update is navigational queries. **While this is more a August Core Update signal, I believe it also plays a role in the Helpful Content Update** (more on this later).

There are so many navigational searches for the term that it is now integrated into the Auto-Suggest.

I highly suspect that the Helpful Content Update score is impacted by user data such as "Navigational Queries" and "Returning Visitors".

The logic being that if users are seeking or returning to a website, then it must be helpful.

"Bad" Content Provides Ranking Clues



This is the Ahref organic traffic graph of a **828 word article with no traditional trust signals which still rankings** after the Helpful Content Update for the term: "*top cigar lighters*".

Here's how it looks:

famous SMOKE SHOP Since 1939 Best Cigars Selection Guide

Search

Home » Best Cigar Accessories » Top 10 Cigar Lighters

Select a Category

Top 10 Cigar Lighters

Lighters give you a quick and easy solution to lighting cigars and keeping them lit too! With many different options, colors, and formats, we've condensed them into a list that we think showcases the best lighters available!

Xikar HP3 Triple Flame

The Xikar HP3 Triple Flame lighter is the ultimate in High Performance. Its triple-jet flame is complemented by an oversize flame adjustment knob that allows you to dial in the perfect flame height. As always, Xikar's lifetime warranty is standard.

Buy Now

Best Cigar Accessories

- Top 10 Humidors
- Top Cigar Travel Cases
- Top 10 Cigar Cutters
- Top Cigar Ashtrays
- Top Humidification Devices

Related Articles

- 5 Of The Best Cigar Lighters For The Money
- 5 Most Reliable Cigar Lighters
- Best Performing Lighters For Cigars
- Torch Lighters With 3 Flame Jets

This is highly unusual because there are absolutely no traditional "trust" or "helpfulness" signals on the page.

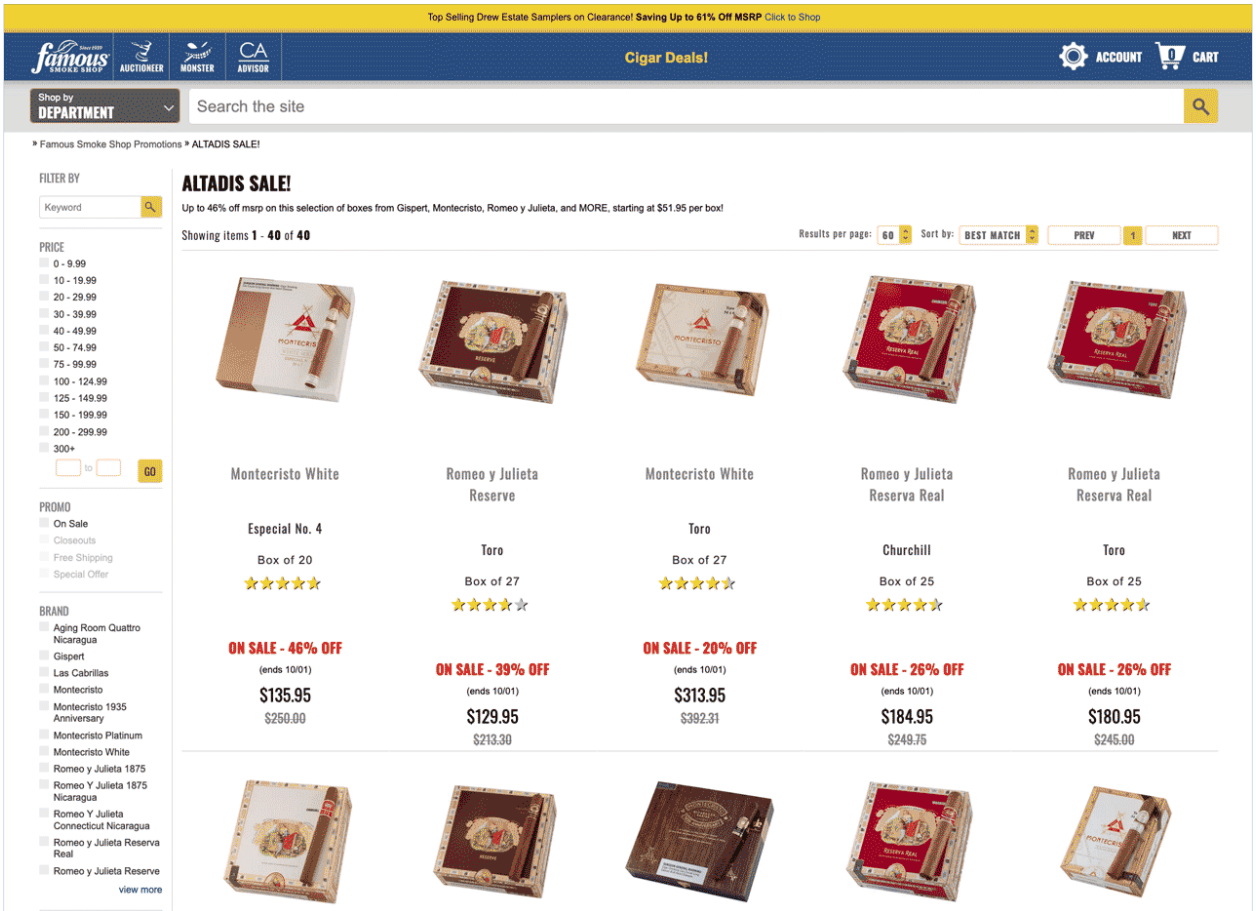
- There are no authors.
- There is no publish date.
- There are no names mentioned anywhere.
- There are no original images.
- There is no proof that the writer used any of the products.

In fact, this isn't really an article... it's almost like a eCommerce product listing or a resource page.

And that might be the point.

This website IS an eCommerce and actually sells these products.

While nearly all pages on this portion of the site would likely fail the Helpful Content AI checks... **the rest of the site is likely passing by virtue of not qualifying for the verification check.**

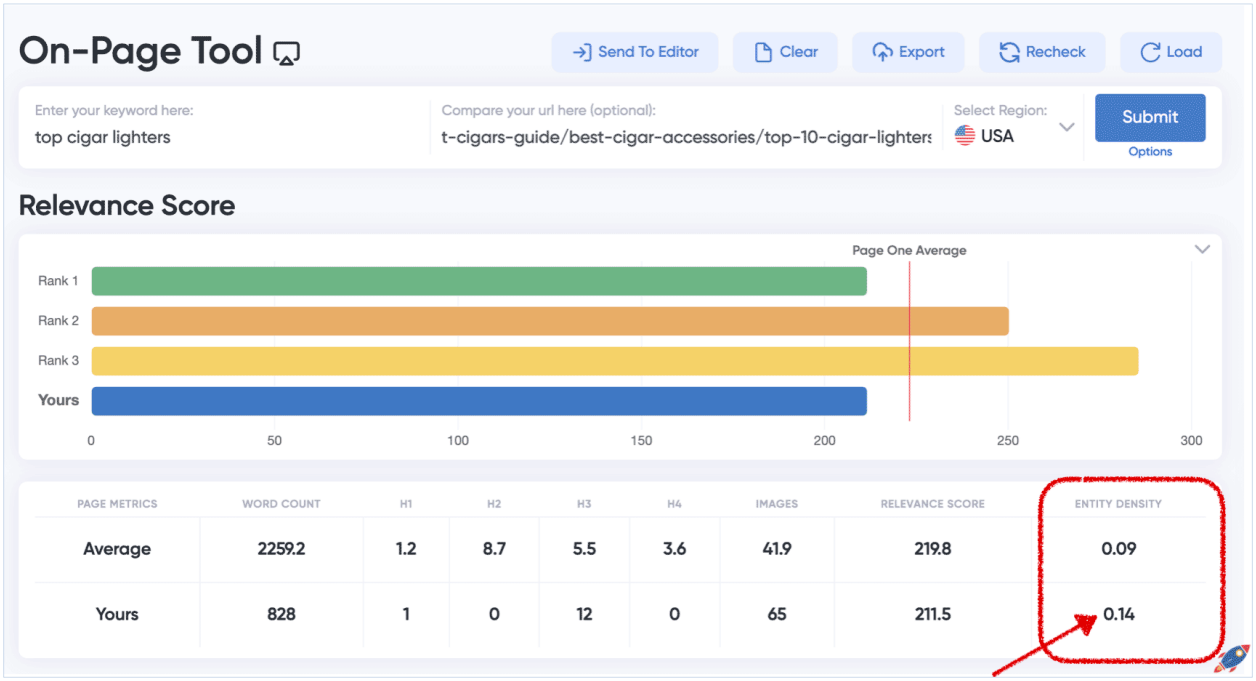


The website has thousands of pages of **eCommerce products that likely outweigh the relatively few article style articles also hosted on the same domain.**

The Helpful Content Score is a site wide score which means that if most of your pages are deemed helpful, then your entire site will benefit.

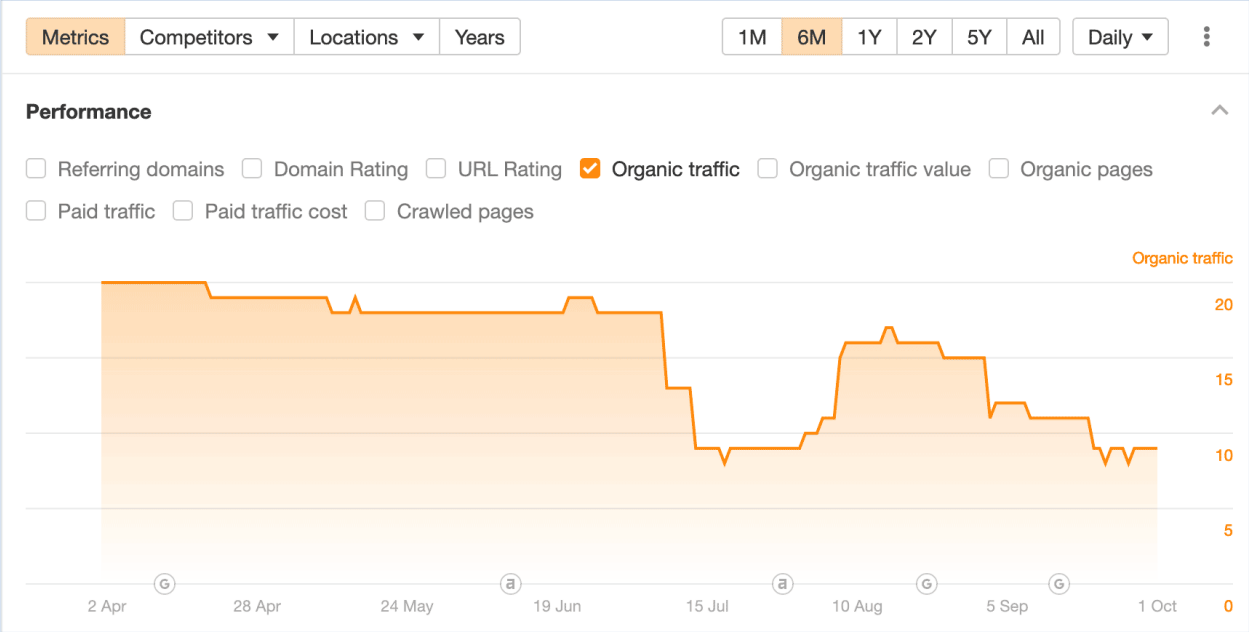
In a podcast between Glenn Gabe and Paul Vera, it was mentioned that Google is looking into rolling out a more granular version of the Helpful Content Update which might address "parasite SEO" by targeting sub-directories and different zones within a website.

It's entirely possible that in the future, Google will be able to distinguish helpful sections of a site versus unhelpful sections... but for now, this is proof that it is indeed a site wide signal.



It appears as if once the article is deemed to be on a helpful site, then Google falls back to favoring entity density to as a primary ranking factor.

No Rewards For Real User Pictures



This is the Ahref organic traffic graph a website that I believe the Helpful Content Update should have rewarded...

And yet, it remained steady while some articles even declined a little. I stumbled upon this website by accident and **it is one of the most genuinely helpful and 'real' website in terms of user experience.**

Here's how it looks:

USA

7 Indian Restaurants You Must Visit in South Florida

Author **David** Date **December 18th, 2019**

South Florida is a food-lovers' dream. Over the last several years, the food culture here has exploded. More and more options are added every time you turn around. But while South Florida, and Miami in particular, is mostly known as a hotspot for Latin cuisine, it also has lots of Indian restaurants you must visit in South Florida.

In South Florida, you can, of course, try dishes that have become popular in the West like butter chicken. But you'll also find lesser-known dishes in the States like marbar beef, bhujia, gujar halwa, and laal maas. You can find many of them in the Miami area. It makes this tropical paradise the perfect place for an Indian food fanatic like me! When you visit South Florida, you should check out the Indian food scene here, too. These are the 7 Indian restaurants you must visit in South Florida!

Bollywood Masala – South Miami

Located in an unassuming strip mall at 57th Avenue and 74th Street in South Miami, Bollywood Masala is a hidden gem in South Florida. This small establishment is big on authentic Indian flavors and has become a favorite of mine and my family! It's located just blocks from my home and I pay them a visit at least once a week!

Chapters

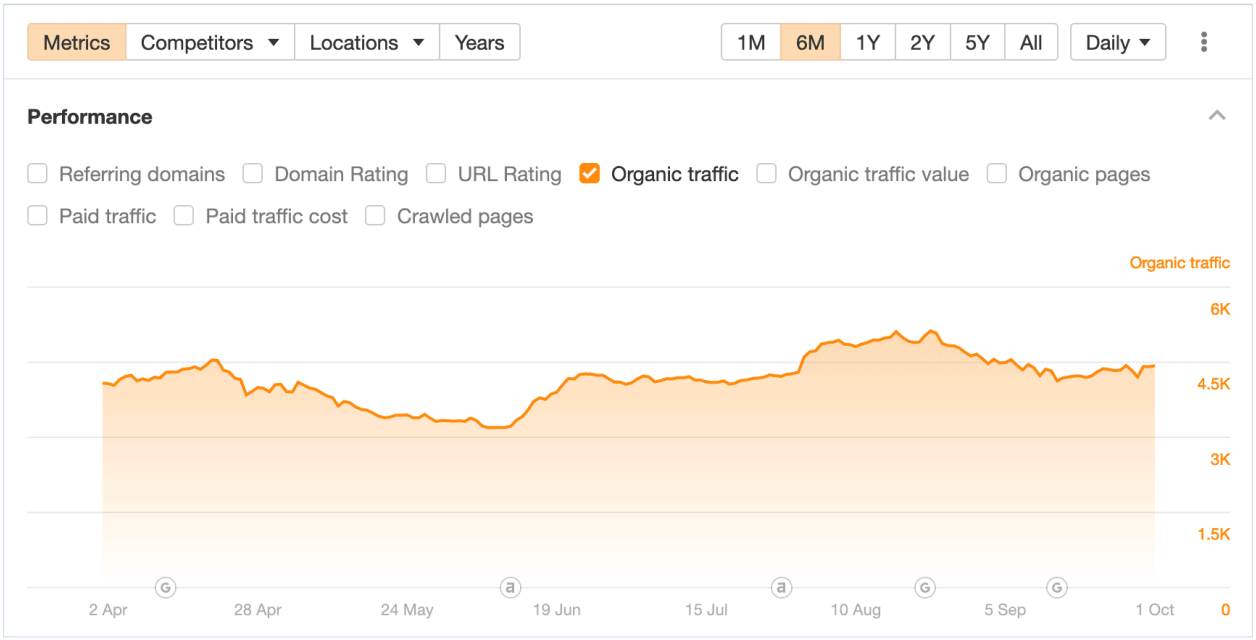
This website owner literally travels around while taking pictures and posts about it on his website. Every-single-picture is taken by him and every recommendation is a place that he's actually been to!

This is *actually* helpful content.

Yet, the Helpful Content AI classifier does not recognize it. Or at least, if it does, it does not reward it. This, along with many other helpful websites I spotted throughout my research, **make me believe that the Helpful Content Update is only a penalty and has no ability to "reward" a website, in spite of what the Google documentation says.**

This is bad for the website owner but great for learning more about how the AI classifier works.

For reference, the rest of his site remained stable throughout the Helpful Content Update.



Here are some takeaways:

1. Original Images

While Google has made a big deal about original images within their content guidelines, in this case, it appears as if they are not being explicitly rewarded.

2. Helpful Content Update might only penalize websites

It is possible that the websites and pages that have seen increases in rankings during the helpful content update have done so through the adjustment of "other" factors such as returning visitors or navigational queries which have seen a surge since the August Core Update.

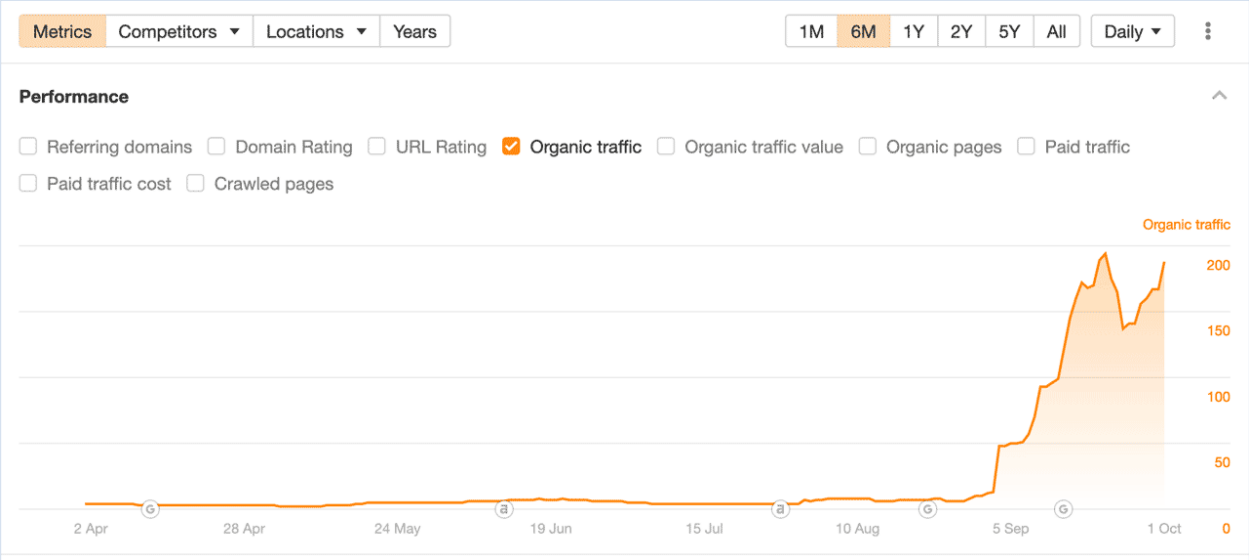
In the absence of returning visitors and navigational queries, it's plausible that the best thing a website could do after the Helpful Content Update is remain stable.

3. Perhaps "Trust" Signals are not being picked up by the Helpful Content AI Classifier

In this specific example, the simple author name "David" is not clickable and does not lead to an author page. The page does not include any author information and just information about restaurants he's visited.

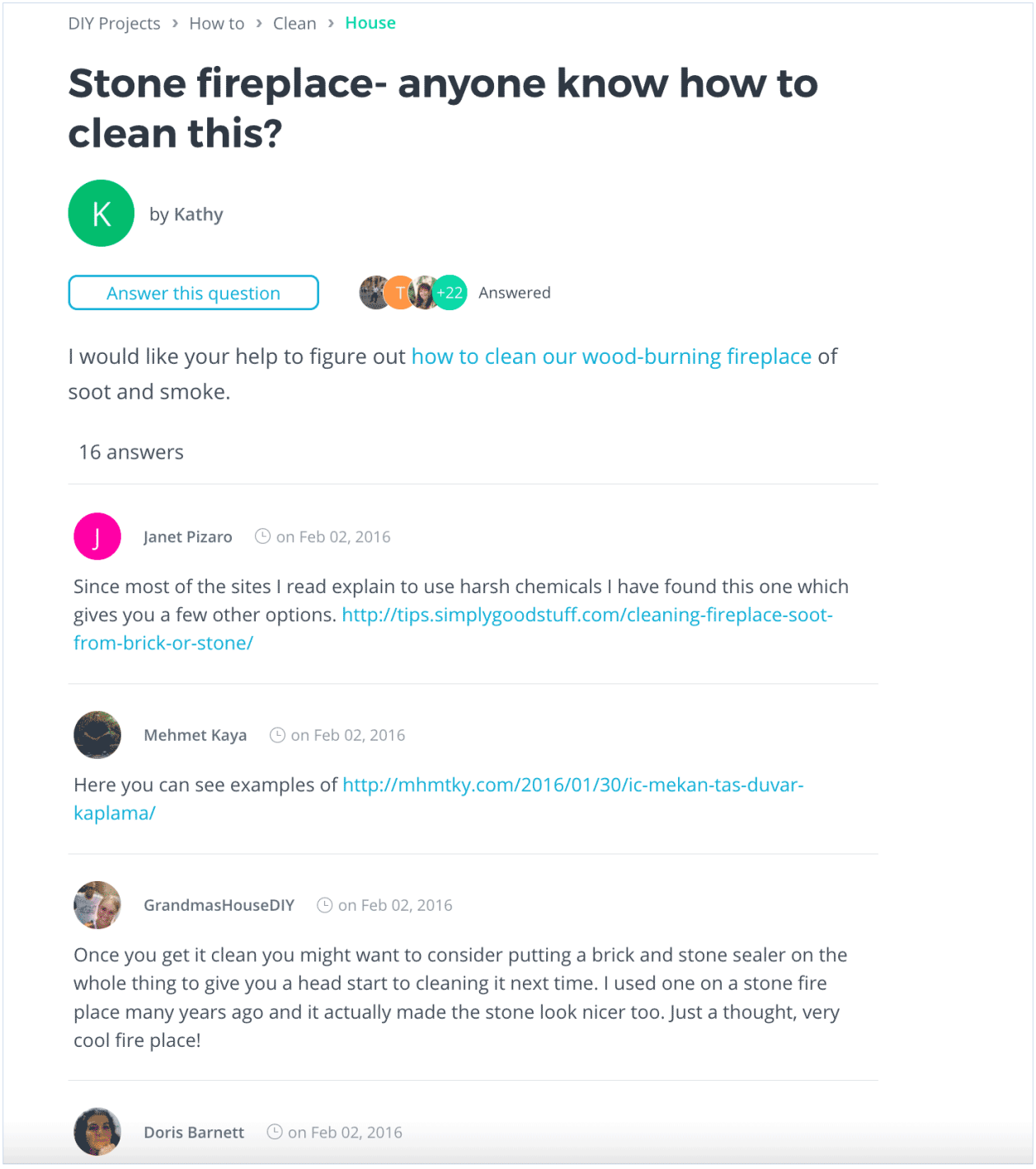
It's entirely possible that the AI classifier cannot locate any trust signals as it cannot process the images to understand the page.

Cleaning Query Offers Solid Clues



This is the Ahref organic traffic graph a website that has been **increasing in traffic after the August Core Update and continued to increase throughout the Helpful Content Update.** I believe this website helps reinforce current theories about the Helpful Content Update.

Here's how it looks:



This website isn't actually helpful. In fact, it's a collection of comments from people that don't really have a good answer to the original question of "*How to clean a stone fireplace*".

There is no guide.

There are no solid solutions.

Just random comments from unqualified users.

Yet it ranks! This is GREAT for figuring out what the Google algorithm is actually rewarding.

It appears as if this is ranking for one of two reasons.

Either

A) **The website has a high quantity of direct and returning visitors.** The proof is that all these people have profiles and are commenting on this thread.

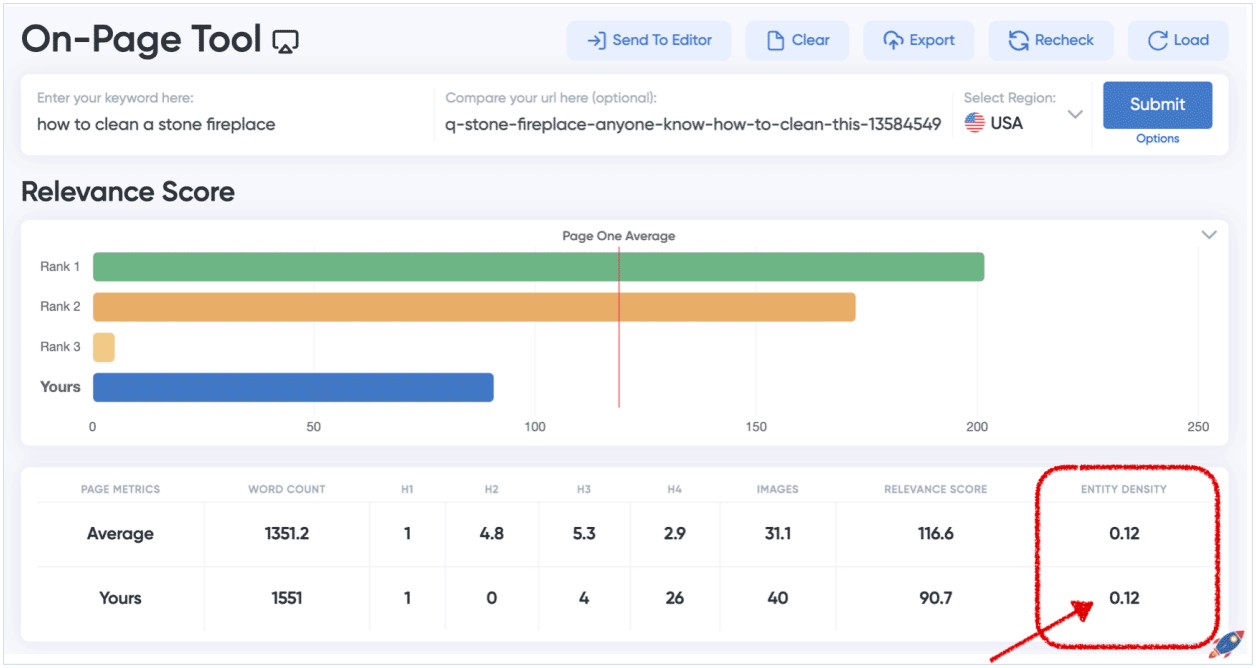
OR

B) **Google is heavily favoring names and usernames with identifiable profiles.**

As far as I can tell, there are no other reasons why this content should be ranking. There are no backlinks, the content isn't special and as shown below, the entity density is equal to the average. No more, no less.

C) **The AI Classifier only detects potential answers and thinks this page is hyper-focused on the topic.**

While none of the answers are actually good... the AI classifier might not realize this and might just interpret this page as exclusively having answers to the question. This could, incorrectly, lead it to assume the narrowly focused page is a great result for the query.



I personally suspect that this is an indication that the larger sites with returning visitors are impervious to being penalized and maintaining their advantage while other sites drop during the Helpful Content Update.

Due to the large quantity of returning visitors and navigational queries caused by community discussions, we see a correlation between multiple usernames / names and high returning visitors.

This poses a challenge when trying to identify specifically which one Google is rewarding.

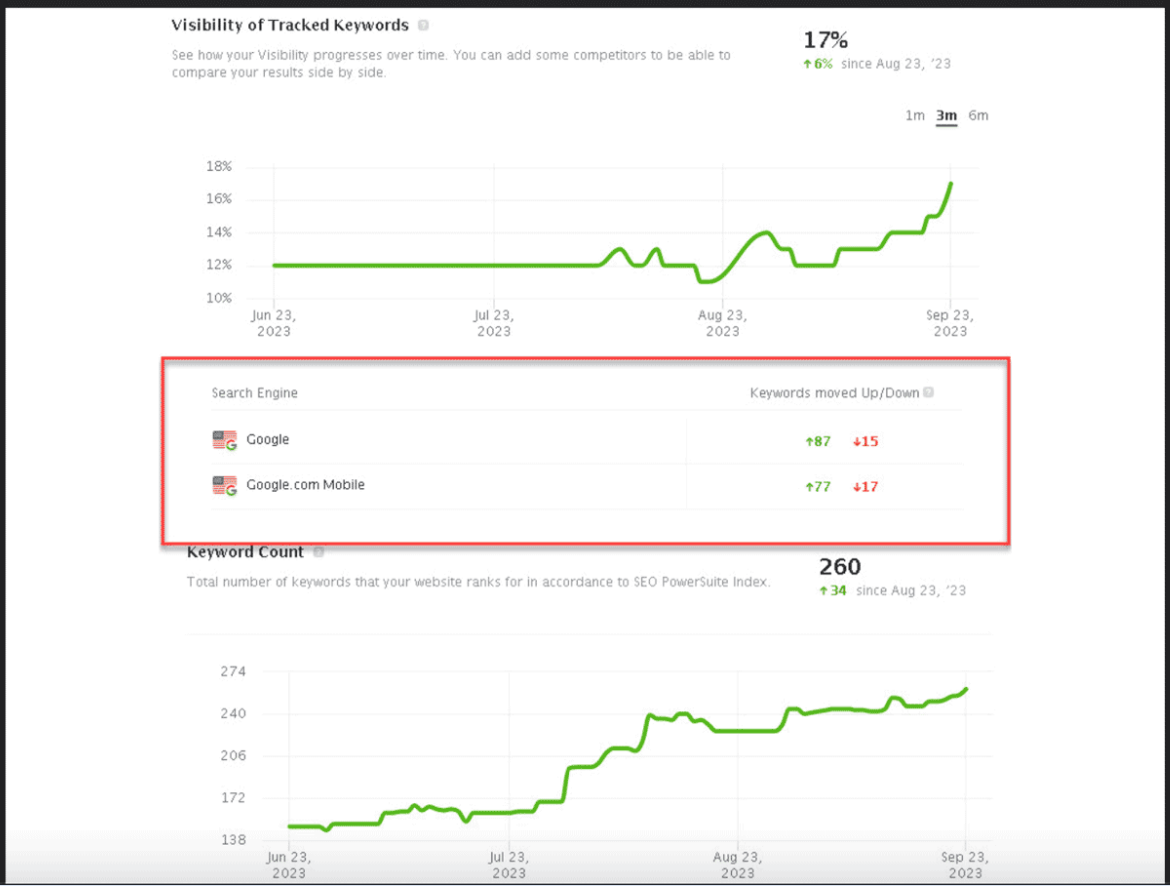
Anecdotally, it seems like users of CTR manipulation Facebook groups are seeing success however please take this with a grain of salt as they have a heavy incentive to claim that click-through-rate (and queries for their terms) seems to be helping.

Again, the ONLY thing that has changed on this site is CTR. No additional links, nothing. It is performing BETTER than a similar site that we have had a VERY well-known SEO work on for us, using PR, backlinks, cloud stacks, etc., over the same period. We have actually started CTR again that site now as well, and it's lifting, too!!

And just so you know, it's not all puppy dogs, sunshine, and butterflies; here is another site (the last image). We have been running CTR against it as well, yep, down. 🤢 What am I going to do about it? NOTHING! Because I TRUST THE PROCESS!!

You guys who keep yo-yo-ing around are like those guys who jump from lane to lane on the highway and end up WAY behind what they would have been if they'd just picked a lane and stuck with it.

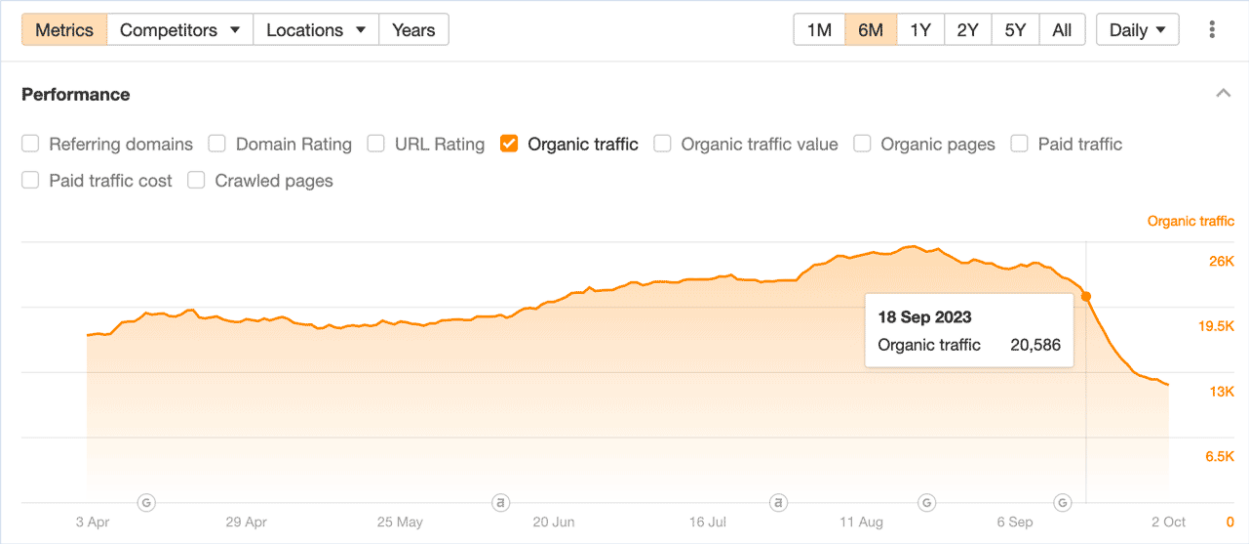
Educate yourself, Implement, and TRUST THE PROCESS!!



Please understand that I am in no way encouraging people to attempt to manipulate click through rate or even send fake navigational queries to Google. Do not manipulate clicks to Google. **I do not believe this is a viable long term strategy** because eventually, when the fake signals stop, you risk seeing a very sharp decline.

I do believe, however, that this **might be an indication that navigation queries are doing quite well post-August Core & Helpful Content update.** Creating a website that encourages visitors to return might be beneficial.

Helpful Home DIY Is Penalized Due To Missing "Signals"



This is the Ahref organic traffic graph a great home DIY website that has been penalized by the Helpful Content Update. I subjectively believe this is a perfect example of the algorithm "getting it wrong" as this is a genuinely helpful website.

(*Note: I do not know the website owners and I am not associated with this website in any way. It was provided to me by my research assistant and upon review, I believe it's a helpful website that just isn't providing the correct signals to Google.)*

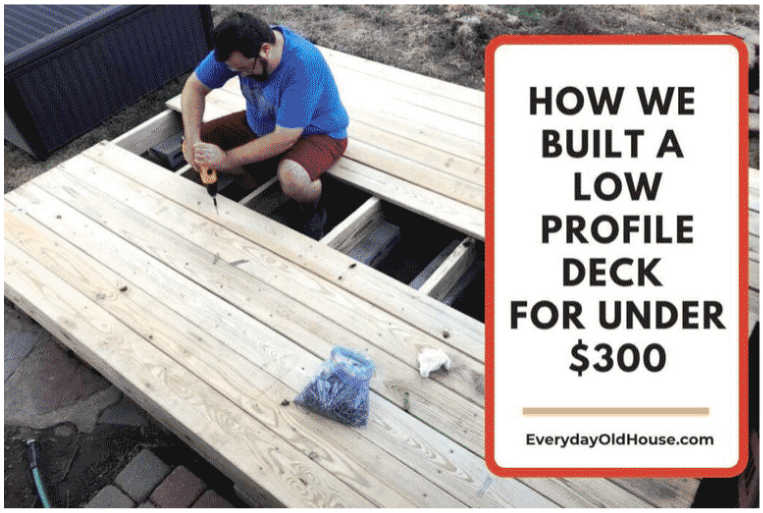
Here's how it looks:

EASY DIY

How We Built a Low Deck for $300

by Jennifer Osterhout | May 10, 2023

Building a low deck is an easy and inexpensive way to add functional space to your backyard. And who doesn't love a bit more outdoor space for entertaining and relaxing with friends and family? Here's how we built a low deck (also known as floating deck) for around $300.

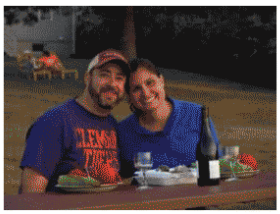


This post contains affiliate links, including but not limited to Amazon Associates. As such, I earn from qualifying purchases. Full disclosure located here.

Backstory

There used to be a hideous spot in our backyard that desperately needed some TLC. It was about an 8 ft by 8 ft square adjacent to our older raised deck.

It's so hideous that not barely even weeds grew there.



WELCOME!

Owning a home where you can unwind and relax can be one of life's most rewarding experiences. But homeownership can also be challenging and stressful, especially with old houses.

Hi! I'm Jen, an "everyday" homeowner on a mission to make homeownership a little easier and a lot more fun. Care to join me? Read on….

LET'S FOLLOW EACH OTHER! (BUT NOT IN A CREEPY WAY😉)

Subscribing is the best way to get updates and snag useful home tips!

Newsletters are usually once a month (maybe more if I ever get my act together…). NO spam, I promise! I hate spam….despise it as much as 1970s vinyl floors. Yuck….

These home owners perform and meticulously document their DIY adventures with detailed steps, receipts, photographs and explain how to replicate their results.

This is, in my opinion, the holy grail of "first hand experience".

Yet, they were severely penalized by the helpful content update.

This provides us with a great opportunity to identify "Why" the AI classifier might disagree with a human opinion. (*Of course, I'm sure other humans might disagree with my assessment, claiming that if it's not an experienced carpenter, then this should not be allowed. However I personally found it helpful as they outline every part of their budget, all the materials they purchased, explain the pitfalls and document each step of the way with photos. This is exactly the guide I would be looking for if I was looking to build a cheap low deck.*)

So what can we observe?

1. There is a single author name however upon clicking it, there is no bio or additional information. There is next-to-no information about the author.

While there is information about the author on the sidebar, this information would likely never be used by Google's AI classifier because the sidebar loads AFTER the main content within the source code.

From a computer's point of view, this is equivalent to the the sidebar being located below the main content which exceeds 3000 words. If my theory with the AI classifier being limited to approximately 2000 words, it would mean that the AI classifier never gets to the 'expert information' part. However...

2. Even if the AI Classifier did read the bio, it might not deem the authors as experts.

The poorly written bio provides no clues as to why they are trustworthy. While it might be obvious to humans that the authors have first hand experience with their projects, the AI classifier does NOT see/understand the images and might be seeking mentions of a contractor, wood worker or something similar within the bio.

3. The AI learns by processing patterns over and it might deem pages that contain a prominent disclaimer such as *"this post contains affiliate links..."* as potentially less trustworthy.

I do not think this is the case however if the AI was manually fed thousands of 'bad' examples that contained a similar string, then it could react that way.

Ultimately, this site was impacted in spite of having dozens of original, hand captured photographs documenting every project they perform. **I believe this should be enough evidence that original photography won't save you from the Helpful Content Update.** In spite of that, I am still a big proponent of original photography.

I also believe that we are seeing the AI classifier look for 'trust signals' within the text. When you imagine that this is what the AI classifier is processing:

```
How We Built a Low Deck for $300
by Jennifer OsterhoutMay 10, 2023
Building a low deck is an easy and inexpensive way to add functional space to your backyard. And who doesn't love a
bit more outdoor space for entertaining and relaxing with friends and family?  Here's how we built a low deck (also
known as floating deck) for around $300.
This post contains affiliate links, including but not limited to Amazon Associates. As such, I earn from qualifying
purchases. Full disclosure located here.
Backstory
There used to be a hideous spot in our backyard that desperately needed some TLC.  It was about an 8 ft by 8 ft square
adjacent to our older raised deck.
It's so hideous that not barely even weeds grew there.
After some research (aka scrolling through Google, Pinterest and YouTube) we found the perfect solution to transform
this wasted space into a space where our family could congregate.
What did we do? We built a low, otherwise called, floating deck.
What is a Low or Floating Deck?
A low deck is called many things, including a "floating" or "freestanding deck".  Whatever you want to call it,
basically it is a deck constructed in a way that it appears to "float" on top of the ground.
And no, it's not magic and really floating. The wooden deck portion sits on a series of concrete blocks that most
folks usually can't see, making the deck appear to be "floating".
Photo courtesy of The Fixer
Pretty clever, eh?
In addition, since it's low to the ground, these decks don't need handrails or stairs.  This keeps building a low deck
not only easier for DIYers, but also keeps material costs down.
Do You Need a Permit for Building a Low Deck?
Constructing traditional decks attached to your home usually requires building permits.  However, low decks usually
don't require a permit as long as they:
        •       Are truly "low" (i.e. less than 12 inches above the ground)
        •       Covers only a small portion of your yard
BUT please check with your local building department to confirm!  I'm NOT an expert in this area!
Also – if you plan to dig more than a few inches of soil to build a low deck over grass, consider calling the Dig Safe
hotline (8-1-1 or 1-800-DIG-SAFE).  This is a free service that will come to your home and mark underground utilities
so that you don't have the uncomfortable experience of hitting something.
Supplies for Building a Low Deck
Here's of list of tools (aka what you probably already have in your toolbox or work area) and materials (aka what you
will probably need to buy) we used to build our low deck.
Tools
        •       Spade, Rake and/or Shovel
        •       Tamper
        •       Pencil
        •       Tape measure
        •       Level
        •       String
        •       Framing Square
        •       Circular Saw
        •       Drill
Materials
Note that quantities are for building a 8 ft by 10 ft low deck. Please modify based on your desired deck size.
        •       Spray Paint
        •       21 Concrete Blocks (4in x 8in x 16 in)
        •       Deck screws – 3 1/2-inch and 2 1/2-inch
        •       Seven 8ft long pressure-treated lumber (2in x 6in)
        •       Two 10ft long pressure-treated lumber (2in x 6in)
        •       Seventeen 10ft long pressure-treated 5/4in x 6in deck boards
        •       Landscape Fabric and/or Gravel (optional)
```

It becomes much clearer that the helpfulness signals must fall into this region of text. When the article does not demonstrate 'helpfulness' early on, you are unlikely to be rated as helpful.

In addition, I believe this shows that the **Helpful AI Classifier is mainly seeking 'helpful' signals and doesn't explicitly care about the quality of the writing.**

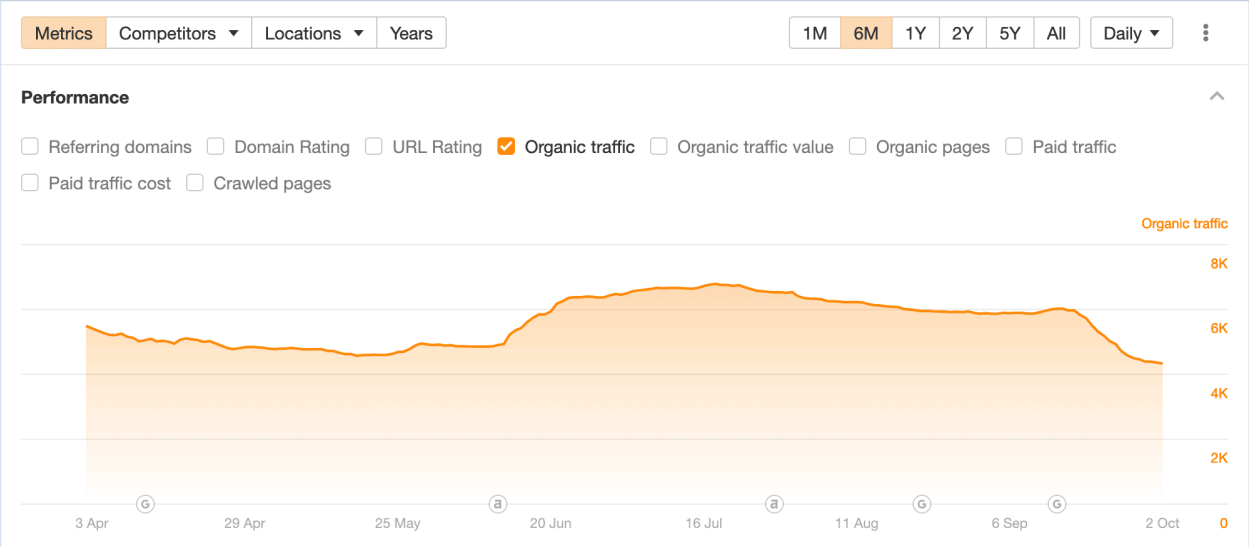
It isn't judging the quality of the material, instead, it's seeking expertise.

(Funny enough, when discussing permits, the author says: "I'm NOT an expert in this area!" I could easily see the AI misinterpreting this and determining that the author is not an expert due to this line.)

According to Google's official recommendations, removing "unhelpful content" from your website can help your helpfulness score increase. If this page is in fact deemed unhelpful, then this would mean that the authors would have to REMOVE the detailed documentation of how they built a $300 low deck in order to stand a chance of recovering. This is absurd as I believe this provides value to the web and it would be a shame to lose this type of real home project.

Fortunately, I have ideas on how to feed the Helpful Content AI classifier what it wants to see.

Helpful Veterinarian Site Is Penalized

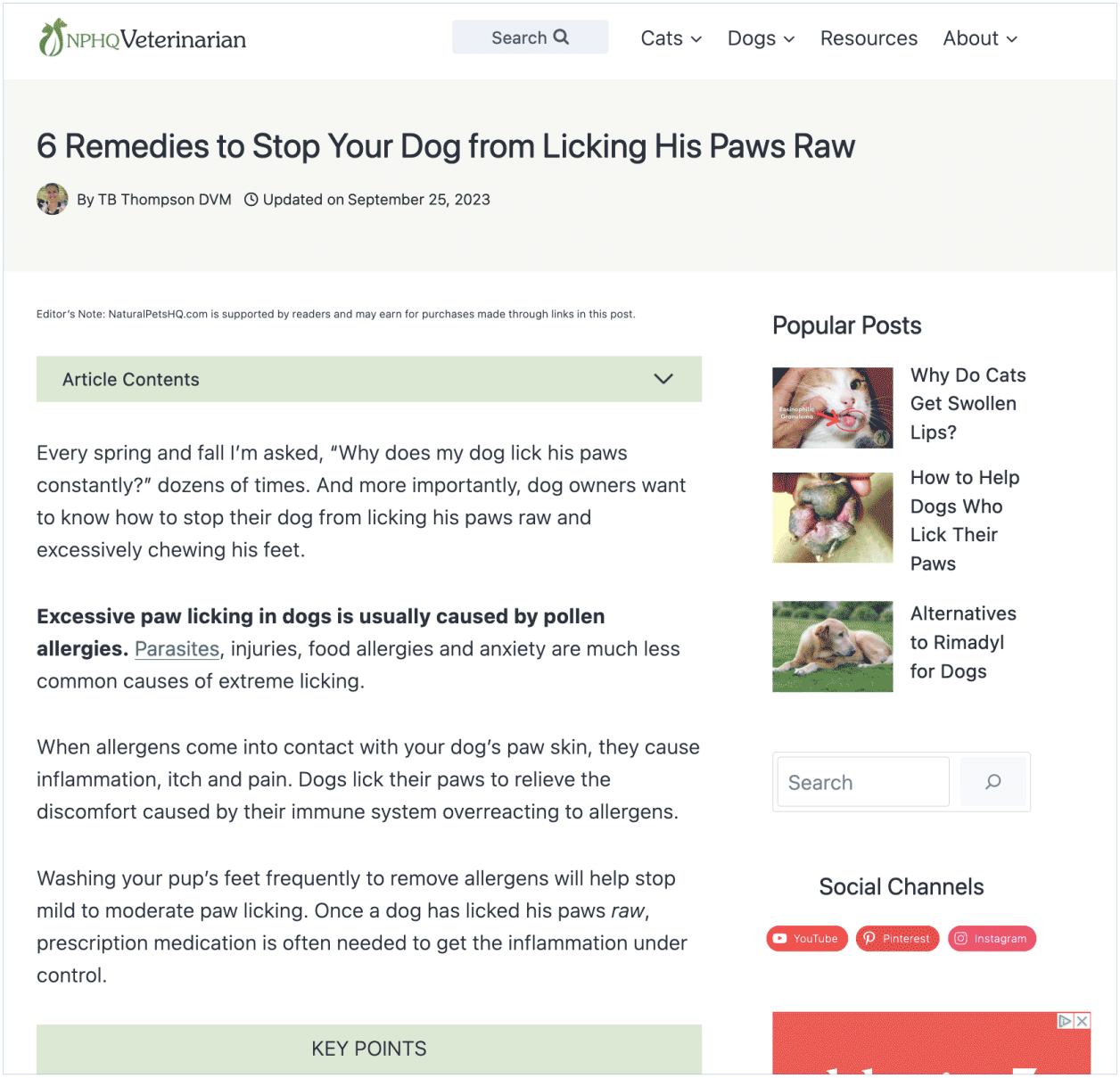


This is the Ahref organic traffic graph a veterinarian website that has been penalized by the Helpful Content Update. This is completely absurd as the information on this website is VERY helpful and there should be no debate as to the expertise of the person writing.

The Helpful Content Update just gets it wrong here.

(Note: I do not know the website owner and I am not associated with this website in any way. It was provided to me by my research assistant and upon review, I believe it's a helpful website that just isn't providing the correct signals to Google.)

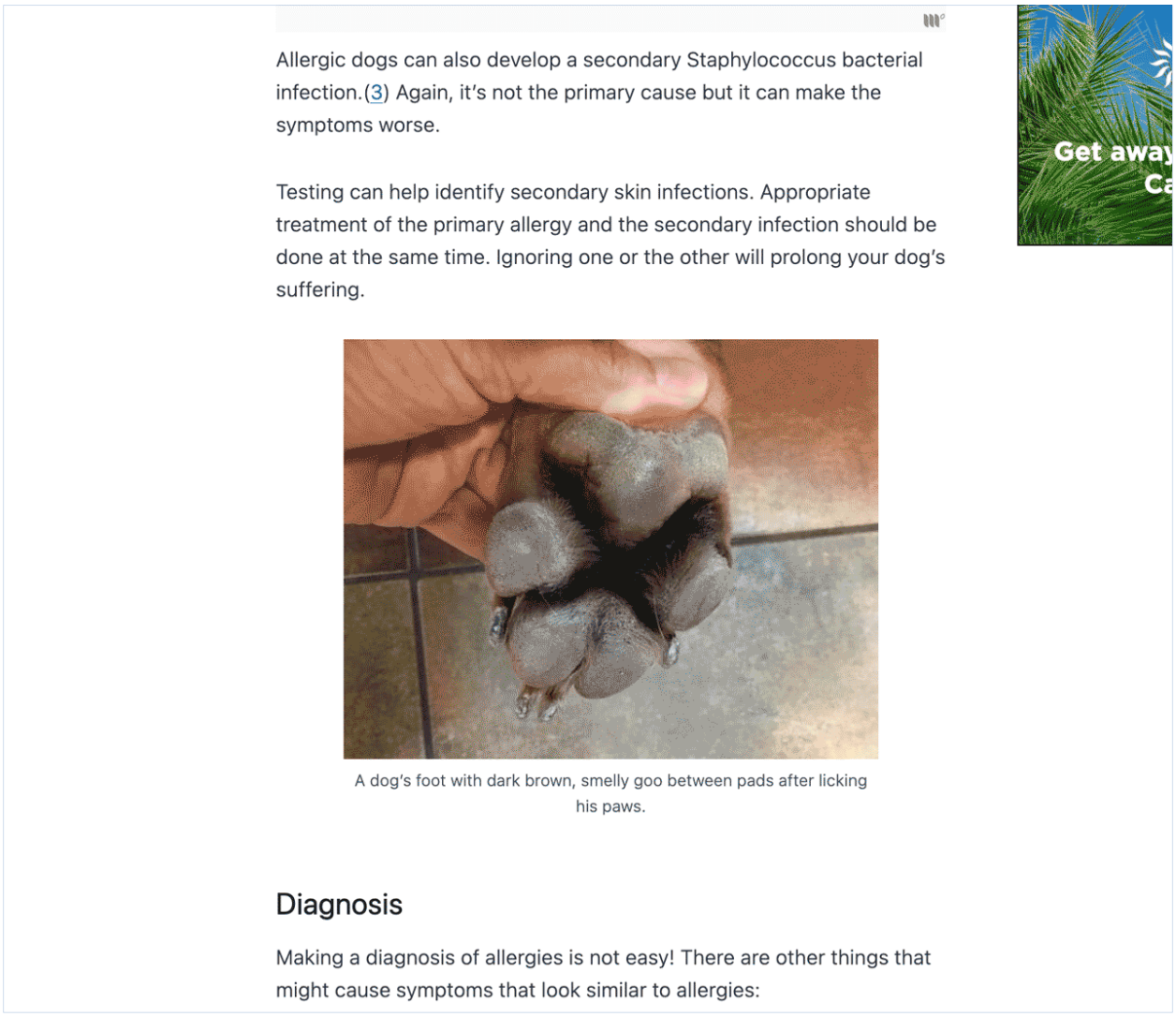
Here's how it looks:



This angers me.

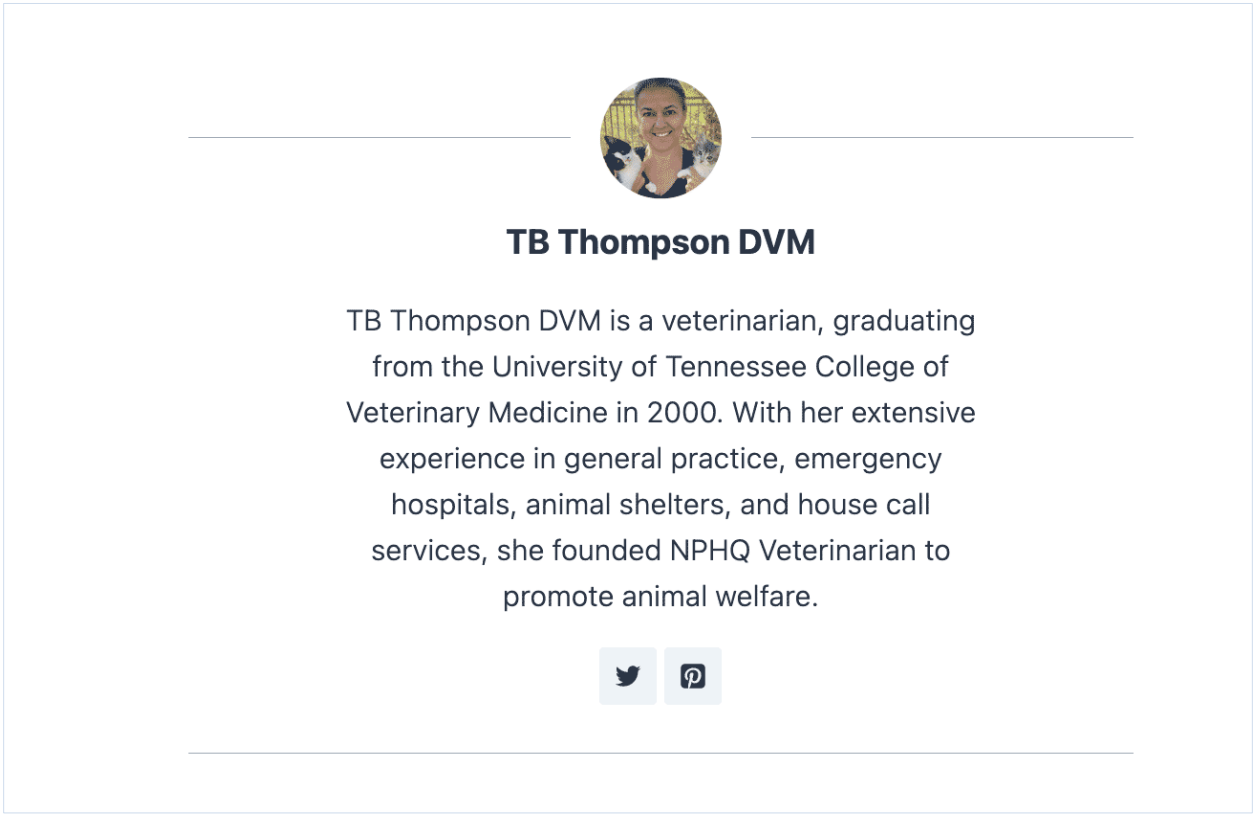
C'mon Google, this is a licensed veterinarian with 20 years experience writing about pet issues. Get off your high horse. This is the type of person I try to help when researching algorithm updates.

Not only does this article provide original imagery:



It ALSO provides medical references throughout the article. The (3) after "*bacterial infection*" points to a medical paper supporting her claim.

And of course, the bio at the bottom.



So what is happening here?

Why is the algorithm penalizing a veterinarian writing about pet health issues, while providing proof of knowledge, expertise, references and original images?

Unfortunately, it's the other stuff.

Putting my SEO hat back on, **it appears as if this article (and the entire website) is overusing display ads throughout the content.** It is excessive and the helpful content update likely deems sites with too many ads within the main content as "unhelpful".

Aloe vera gel

Aloe vera gel (click to see my favorite brand on Amazon) can be applied topically to soothe irritated skin. Small amounts of aloe vera shouldn't cause any problems if your dog licks it. In rare cases, dogs can be sensitive to aloe vera so discontinue its use if the skin starts looking worse after applying it.

Tea rinses

Tea rinses are a traditional remedy for itchy skin. Use dried calendula, chamomile, green tea, nettle, or yellow dock to make the herbal tea.

Add one tablespoon herb to one cup of boiling water, steep for five minutes. Strain off the herb, then allow the tea to completely cool. Fill a dish with the tea and have your dog stand in it for five to ten minutes twice a day.

Alternatively, use a clean cloth soaked with the tea and apply it as a compress to the affected area of skin twice a day.

Antihistamines

In addition, the author is **writing about the complete topic rather than narrowly focusing on the exact topic at hand.** This *might* play into the Helpful Content Update that seems to be rewarding more narrowly focused content.

And unfortunately, upon investigating the sitemap of the website, I discovered there were **many empty or near empty category pages which might lead to thin content being detected elsewhere on the site.**

The page also has an affiliate link disclaimer which the AI classifier might be picking up and classifying as "likely unhelpful".

Finally, **the articles are so long that the AI classifier likely never gets to the bottom of the page in order to read the author's full bio.**

The silver lining is that websites like these help us understand more what the AI classifier is seeking. While the information is helpful *(I accidentally learned something new that I will be applying in my day to day personal life while reading her website)*, the Google helpful update says it isn't.

7. Analysis and Trends

(Here's what keeps on coming back)

After reviewing the accumulated metric data (average word count, sub-headline breakdown, entities, etc) and hundreds of websites through manual review, some trends begin to emerge.

It is important to note that the **August Core Update has had an impact on the recent changes in the search engine landscape and that these trends are building on TOP of what has previously changed**.

I have yet to encounter a website that saw a significant increase during the August Core Update and only to be subsequently penalized by the Helpful Content Update. If you have an example of a site that follows this trends, please send it to me.

Here are some of the observed trends with regards to the recent search landscape:

1. **Shorter, more narrowly focused content seems to rank slightly better.** Overall word count decreased slightly.

2. In spite of content being more narrowly focused, **related entity density remains high** (even higher than before).

3. **AI content does not seem to have been impacted** and still ranks just as well as before.

4. **Content quality (writing style, quality) does not seem to matter.** Poor quality writing still ranks.

5. **Content accuracy does not seem to matter.** While I suspect the helpful content update classifier AI can detect answers... it cannot determine if it is the *correct* answer.

6. **Identifiable answers should be located near the top of the article.** The AI classifier seems to be able to determine if you are providing an answer to the query. It seems to favor multiple answers from various perspectives.

7. **User generated content seems to be favored.** Names (full names and profile usernames) seem to be prominent on trusted sites.

8. **Websites with ample comments below the articles seem to be unfazed by the helpful content update.** This may just be a by-product of having an engaged community.

9. **Mentions of specific expertise within the page seems to help.**

10. Larger websites with **returning visitors do not seem to have been impacted by the update.**

11. Larger websites with **high quantities of navigation queries do not seem to have been impacted by the update.**

12. **Local websites** have seen a minimal / **negligible impact** from the update.

13. **Ecommerce websites** have **not seen a major impact** from the update.

14. **Resource sites** (especially custom coded sites and sites using a CMS other than WordPress) **do not seem to have been impacted** as much by the Helpful Content Update.

15. **The main type of website affected by the update seems to be WordPress** driven blog-style sites publishing full length articles.

16. Having a prominent **affiliate link disclosure** did not guarantee you were impacted by the Helpful Content Update but **seemed to increase the likelihood that you might be.**

17. There was **no observable difference between sites that used stock or even 'no photography' versus those that used original, homemade images.**

18. Most helpful pages contained an **identifiable date** however that date did NOT need to be recent.

19. Sites with **excessive advertisement integrated into the main content seemed to be negatively affected.**

20. Newer sites with a **limited history seem to have been sparred** by the Helpful Content Update.

21. In-depth pages that went into **multiple sub-topics seem to perform less after the update.** (We are seeing less H3, H4 sub-headlines.)

22. **Links (both external and internal) did not seem influence if a website was affected or not.** While I don't go into detail within this analysis, I did spent quite a bit of time analyzing backlinks, domain authority, internal linking structures and so forth. While it is true that major websites (DR80+ seemed to be sparred by the Helpful Content Update, I suspect that this is most likely due to other factors.)

While these are the major trends observed across hundreds of pages, I must restate that correlation does not equate causation. These are just observations across many pages.

8. What I Believe Happened (3 Theories)

Main Theory #1: Helpful Content Update

This is my first of three theories of what happened during the September 2023 Helpful Content Update.

Until we develop a repeatable process for recovering websites, I want to make it very clear that this is just a theory. As the test results from the experiments arrive, I will be able to further support or disprove this theory.

First, **I believe that the August Core Update and the Helpful Content Update are indirectly linked.** While they are two completely different updates, focusing on different portions of the algorithm, I believe that one helps the other.

Here's how:

Running an AI classifier requires immense resources and therefore **it would be beneficial for Google to minimize the quantity of pages that need to be analyzed by an AI.** When working on the scale of the entire web, a 10% reduction in pages that need to be analyzed could result:

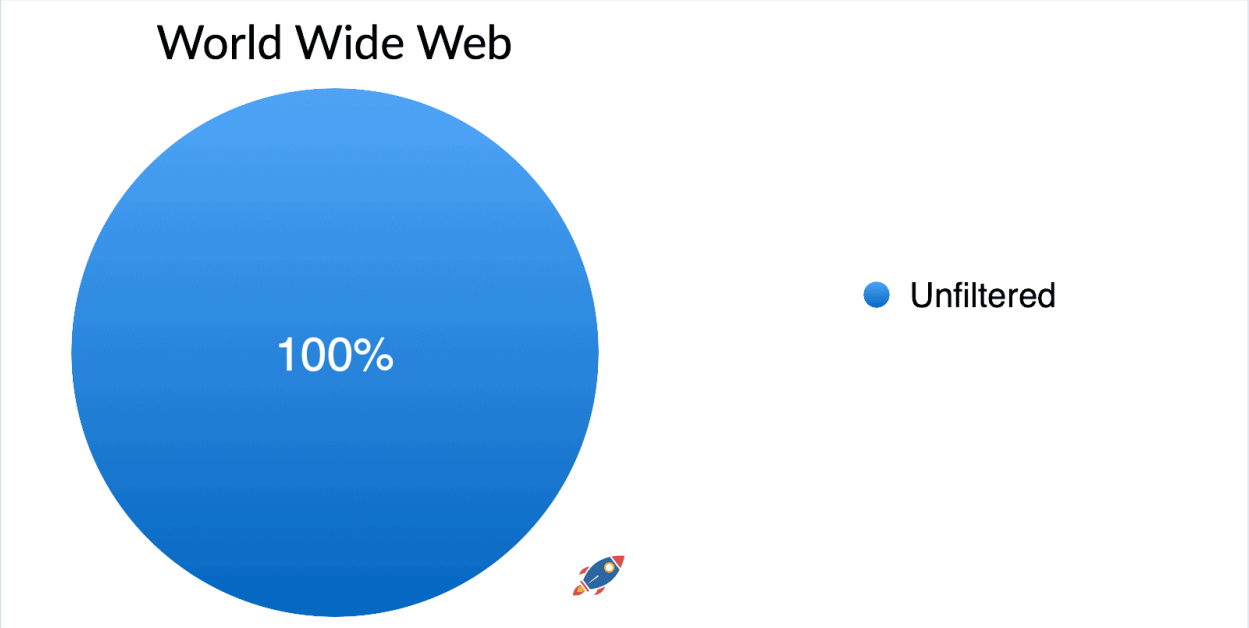
- Billions of pages less to analyze,
- Faster turnaround times

And more importantly...

- Millions of dollars of resources saved.

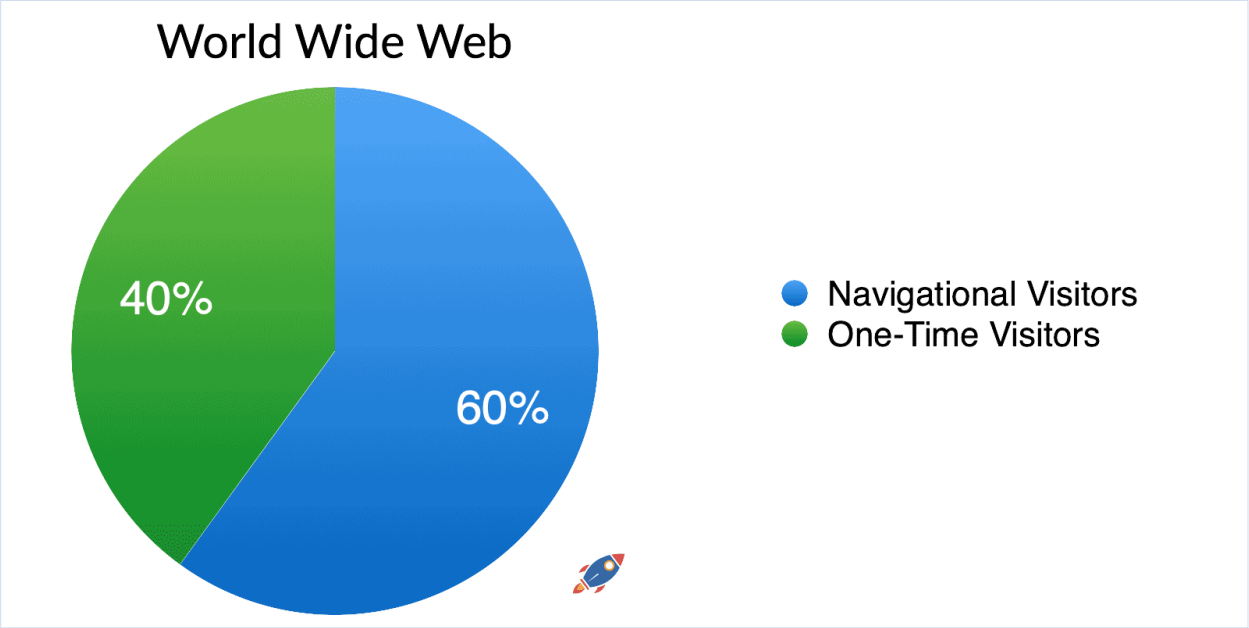
Therefore, it is VERY much in Google's interest to minimize the quantity of pages that need to be crawled by their new Helpful Content AI classifier.

If we imagine the entire web during the August 2023 Core Update.



This is what scanning the entire web would look like if there was no filtering.

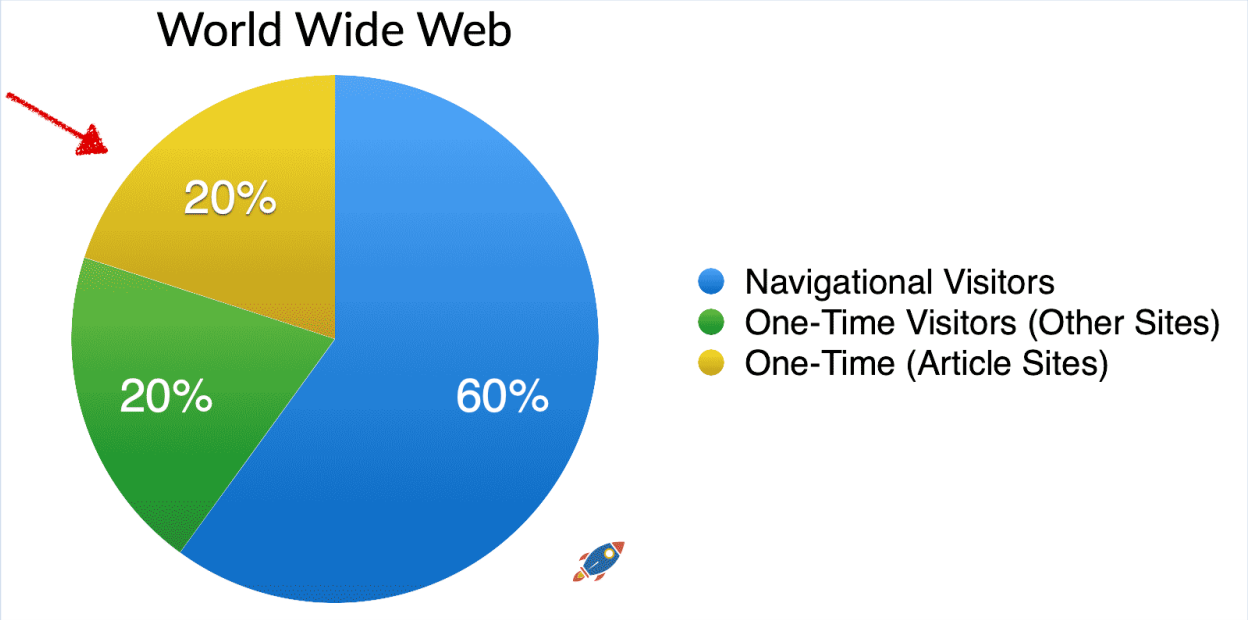
However, Core updates are often focused on metrics so we can imagine Google refining recurring and navigational query metrics. **Specifically isolating websites that might be problematic for further analysis.**



If we work under the assumption that when websites receive many recurring visitors and navigational searches, they MUST be helpful websites, we can eliminate a large portion of the web that needs to be analyzed by the Helpful Content AI Classifier.

This could save Google millions in resources.

And of course, we ALSO know that Google is focused on articles so we can further reduce this number.



Targeting only article sites within the group of sites that do not have sufficient navigational/recurring visitors reduces the quantity of sites that needs to be crawled even further to a hypothetical 20%. (Illustrated above)

By pre-qualifying billions of pages as "Helpful" by virtue of having recurring visitors, Google wouldn't have to process the world's largest websites with their AI classifier.

This could explain the how the August Core Update and the September Helpful Content update work together to process the web in a more efficient manner.

I believe the August Core Update refined and increased the weight of these user metrics (direct visitors / recurring visitors) which is why we initially saw an increase on forums, Reddit and Quora. Not only that, but it would further serve as a 'protection shield' from the Helpful Content Update.

Simply said, if your website has large quantities of users searching for it (navigational) or a high quantity of users returning to it on a regular basis (recurring), then you are automatically considered "Helpful".

I believe this is ONE of the pathways to recovery.

Google has recently published (September 7th) an updated guide on how to set your official site name on Google.

One of the reason it would make sense to encourage websites to set their site name is so they can accurately measure navigational queries.

I believe that one way to recover from the update is to improve your site's metrics by increasing the quantity of navigational queries seeking your website. Alternatively, the other way to recover would be to enhance each page of your website so that the AI Classifier deems the majority of your pages as helpful.

Historically, Google has determined your site metrics over the period of 1-3 months. If it does not have sufficient data, then it cannot provide an accurate assessment of how your site is performing.

Therefore brand new sites (ie: AI spam sites publishing several thousand pages per day) would not have yet accumulated enough user data to be classified as "one-time visitor" sites... and might have been *temporarily* sparred from the Helpful Content Update.

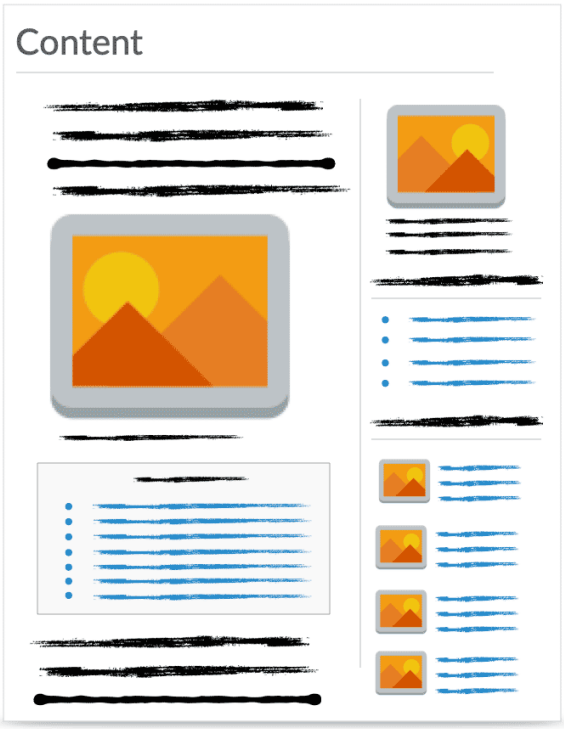
Helpful Content AI Classifier

With respect to the Helpful Content AI classifier, Google has told us is it a model that determines the helpfulness of a page. Ultimately, **I believe the Helpful Content portion of the update is likely ONLY a penalty** and that increases in ranking during the same time frame is likely the result of content being reshuffled (as opposed to being rewarded). In most cases, when we saw significant increases in traffic, they were caused by the August Core Update.

In light of my recent work with AI models, I'm convinced that Google would have to design a very lightweight and resource efficient AI model in order to feasibly process millions of pages.

One of the characteristics of lightweight models is that they are limited by how much content they can process at a time (Model max token input size).

My estimate is that the Helpful Content AI Classifier only crawls the first 2000 words of a page when performing it's analysis.



I believe that within the first 2000 words of a page, you need to demonstrate:

- Identifiable names (authors, usernames, profiles)
- Proof of expertise or first hand experience
- Date (Last Updated or Published)
- Identifiable answer(s)

(And of course, it helps to also be on a topically relevant website.)

And that's pretty much it!

It is fairly easy to feed the AI a body of text and ask it the following questions:

- Can you clearly identify the author(s)? (Yes/No)
- Can you clearly identify a date? (Yes/No)
- Is the author an expert on the topic? (Yes/No)
- Is there proof of expertise or first hand experience? (Yes/No)
- Does the text answer the topic? (Yes/No)

(You can try this for yourself using ChatGPT)

While I'm sure Google's Helpful Content Classifier would be considerably more nuanced, you could imagine how a AI system like that could evaluate the helpfulness of a page.

Site Wide Signal

Google has told us that this is a site wide signal and that removing unhelpful pages can help your overall Helpful Score.

The way this works is that Google slowly crawls the pages on your website and uses a ratio of helpful vs unhelpful pages.

For example, if your website has 100 pages.

- 10 pages aren't valid article pages (homepage, category page, privacy policy, etc)
- 60 pages are deemed unhelpful (FAIL)
- 30 pages are deemed helpful (PASS)

Then you would likely have 30 out of 90 pages as helpful for a score of:

33% Helpful Site Score

Historically, Google has worked with score ranges (like Pagespeed insights) which could dictates how severely you will be penalized.

We can imagine that a site with a 5% Helpful Site Score would see a significant drop in rankings...

A website with a 60% Helpful Site Score might only experience a small drop...

Perhaps anything above 70% Helpful Site Score could be considered good enough and experience absolutely no performance drop. Truth is, unless a Google engineer leaks out the ranges, we won't know.

The only thing we can observe is that some websites were more severely impacted than others, and therefore we can assume that the Helpful Content Update must contain ranges.

This also means that **if you want to recover a website from the Helpful Content Update, you must update enough pages to pass the threshold that Google has set for the Helpful Content Update**. If you have 1000 unhelpful articles buried deep on your website, they could be harming your entire site's performance.

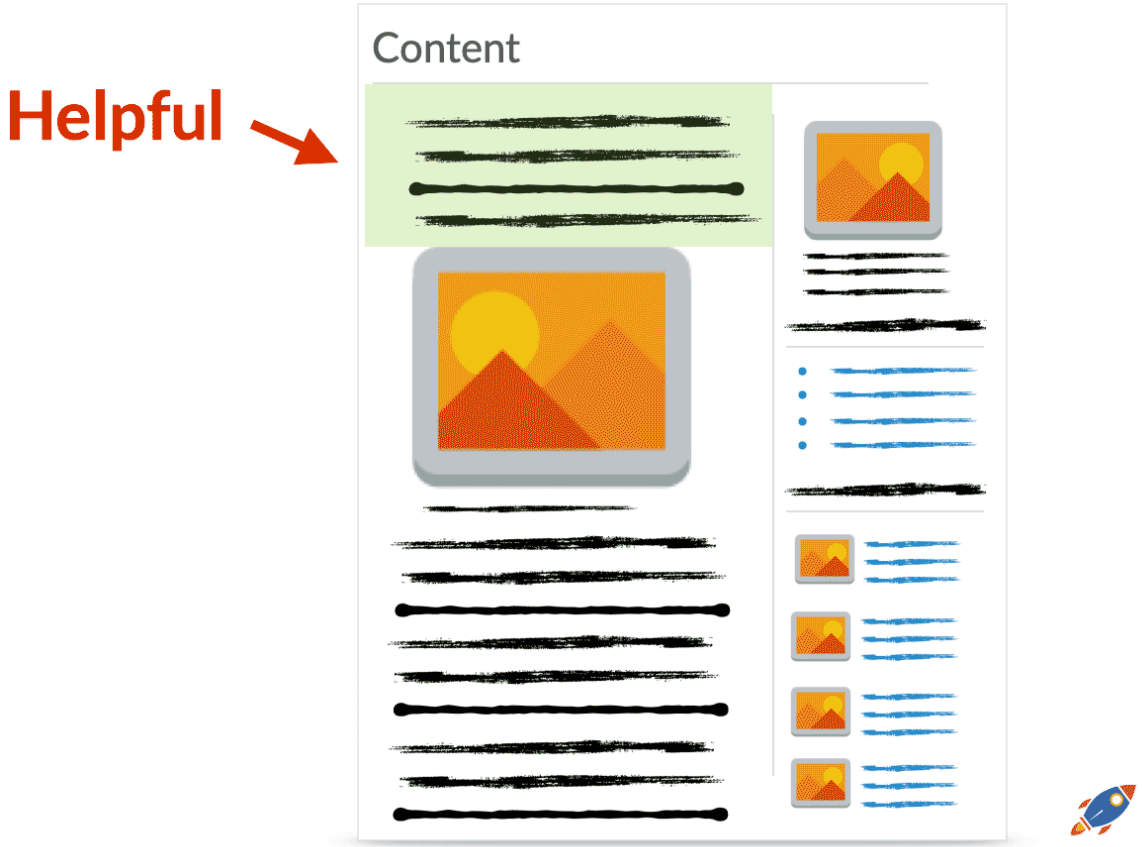
Alternate Theory #2: Narrow vs Wide-Scope Articles

Alternatively, my second theory is that the main focus of this update might be related to the relevance of documents to the search query.

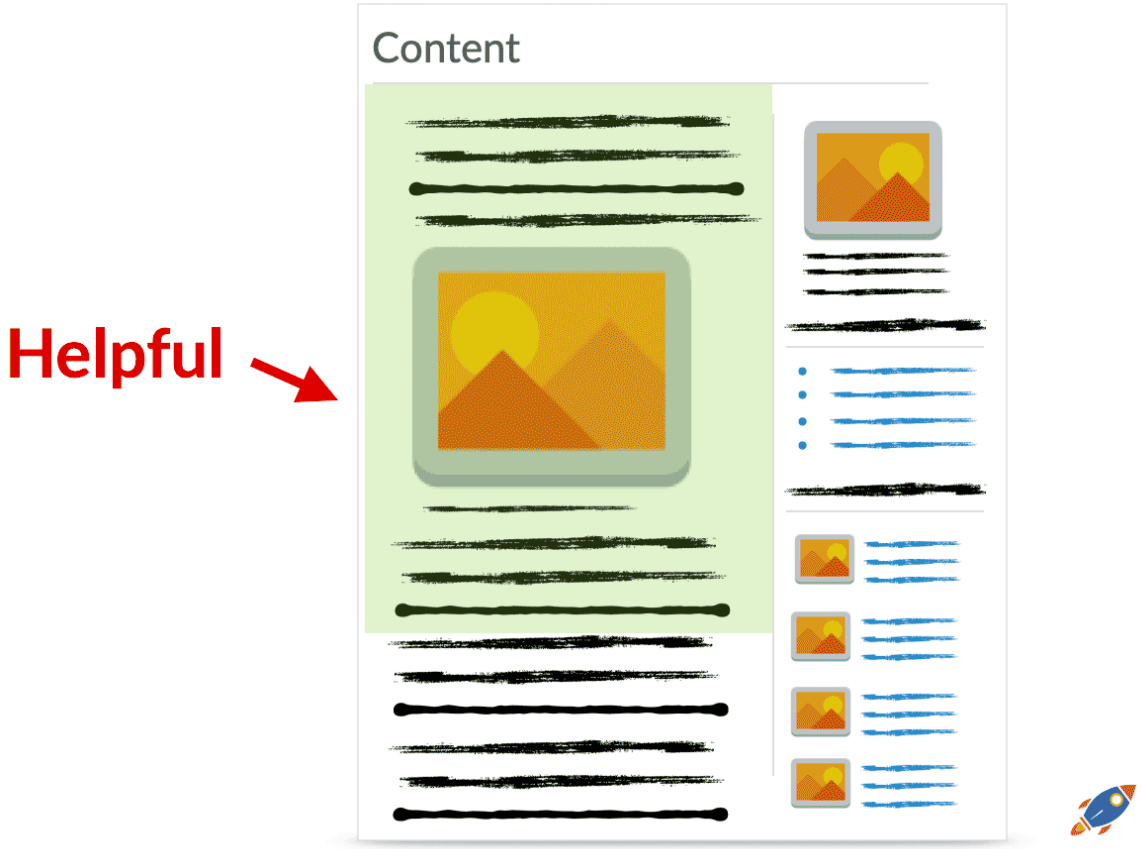
Google's new Helpful AI classifier might be seeking hyper-focused documents to present to users depending on the keyword/title of the page. Data indicates that we are seeing slightly shorter documents ranking overall which might indicate a preference for narrowly focused information.

Hypothetically, an AI classifier could determine the portion of the page that is related to the main topic and determine a "helpfulness score" based on the percentage of a document that relates directly to the query.

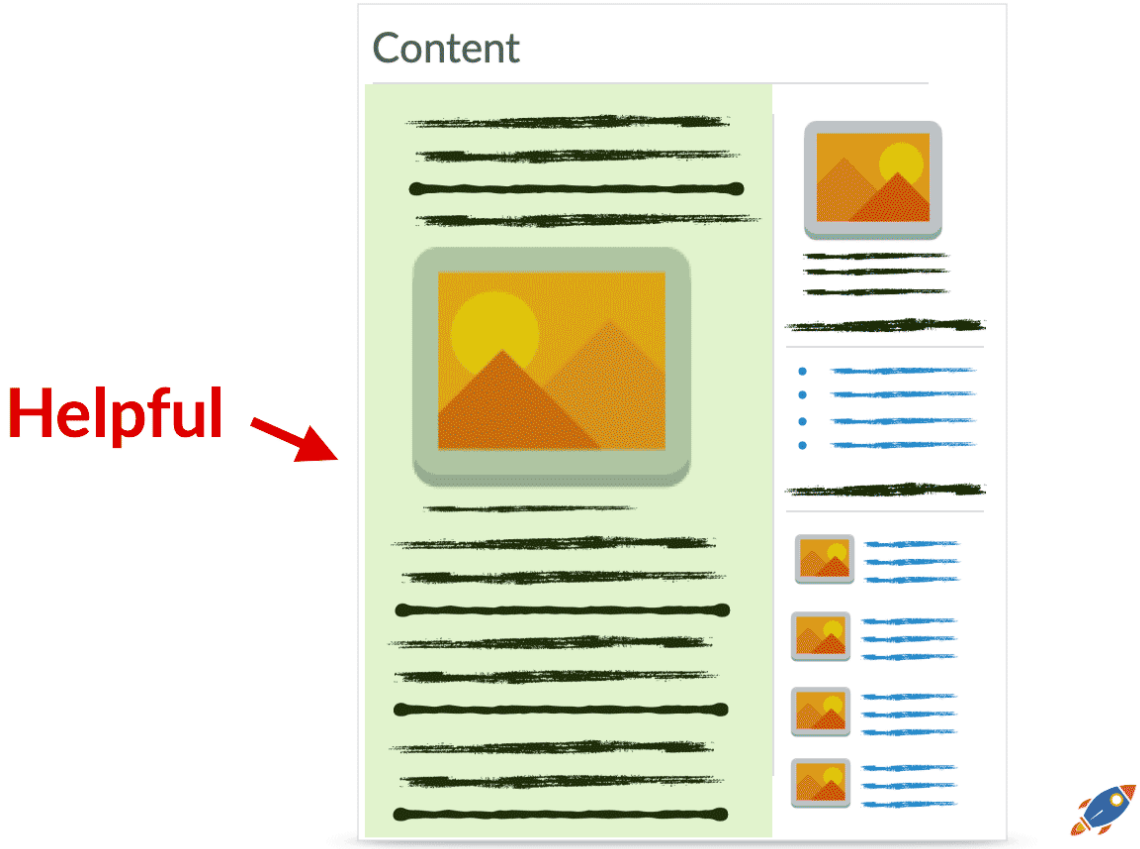
For example:



If only the top portion of the page answers the query, it could hypothetically be determined to be **15% helpful.**



If a larger portion of the page answers the query, it could hypothetically be determined to be **50% helpful.** Passing some sort of arbitrary threshold.



And of course, if the entire page is deemed to be hyper-focused on the query, then the entire page might be deemed **100% helpful.** This could help explain why we are seeing very narrowly focused articles ranking in spite of not meeting the traditional SEO standards.

Google might then, hypothetically, **measure the ratio of articles that feature a wide scope of information versus the narrowly focused articles on our website. And if there are too many "wide scope" articles that don't specifically meet the intention of the query, it would deem that website to be unhelpful.**

For example, if you have a webpage targeting the query: "*how to clean a bathroom floor*" and within that article, you go off topic, discussing the history of bathroom floors, this could potentially be negatively impacted by the helpful content update.

This could tackle the issue of people writing "just for search engines" that Google has been recommending users avoid. Unfortunately, it could also result in very good, long-form content not ranking as well within the search results.

I personally find that academic, in-depth articles on topics can be very useful and often hold the best answers even if that answer only occurs in a very small portion of the article. I selfishly hope that this isn't the approach that Google has taken because I don't think this leads to better content on the web... however as with everything, we'll have to test it.

Alternate Theory #3: Text Classification & Metrics

Finally, a third theory, is that the AI classifier *only* classifies articles into different categories such as:

- Review
- Information
- Resource
- Contact
- Event
- Job posting

...

And upon categorizing all the articles on the web, Google could use traditional user metrics such as returning visitors / navigational queries to determine if the websites hosting information articles are helpful or not.

For example, by categorizing each page first, the AI classifier could subsequently categorize every single website into different categories (ecommerce, informational, local, etc)

After categorizing the web, you could reduce the visibility of every single informational website that is below a certain user metric threshold.

For example:

Step 1) Classify and identify all the information article website
Step 2) If informational website is below a certain user metric threshold, penalize it.

From Google's point of view, this could be the easiest and most straightforward mechanism to re-shuffle the web's search results... but I don't think it would lead to the best quality results. **In this less likely third scenario**, the only way to recover a website would be to either:

A) Change the category of your website by adding a different type of content. For example, if you currently have an informational site that you subsequently transform into a discussion site.

or

B) Improve the user metrics / navigational searches / recurring visitors to pass the threshold set by Google. This could be accomplished by building a community, providing content that is so compelling that users feel they need to bookmark and return to the site or multiple branding campaigns that encourage users to search for your site.

I believe that theory #3 is the least likely theory however I did feel a responsibility to present it as a potential option.

Please note that these are 3 drastically different theories on what happened during the Helpful Content Update. I may retract, edit, improve any theory at anytime as new test information surfaces.

9. On-Going Helpful Content Tests & Experiments

(Validating Theories)

While theories are nice, real world testing is what will determine what really happened during the Helpful Core Update. In order to solidify our understanding of the Helpful Content Update, **we have set up multiple real-world SEO tests to see how Google reacts in certain situations.**

We will be updating this page with the latest test results and encourage you to bookmark this page if you wish to see the updated test results in the future.

Historically, it has taken Google multiple weeks to months to react to changes so we'll have to see how quickly we achieve results.

Test #1 (On-Going)

- We are currently testing improving the author(s) bio, adding specific expertise that lines up with the article topic and making sure that the author information is easily accessible to Google. These are specifically targeted at the Helpful Content AI classifier.

Update: *While Google hasn't had time to fully process every bio author, this has not yet had a significant impact on rankings. I expect to get conclusive results shortly.*

Test #2 (On-Going)

- We are testing specific "Expert Quotes" near the top of articles. These are specifically designed to trigger the Helpful Content AI classifier into thinking there is an answer to the question from an authoritative source. The expert quotes contain names and authoritative answers about the subject at hand. Hundreds of "expert quotes" were injected into content. These should guarantee the Helpful Content AI classifier believes there is a helpful answer from an authoritative source on the page.

Update: *While Google hasn't had time to fully process every quote this has not yet had a significant impact on rankings. I expect to get conclusive results shortly.*

Test #3 (On-Going)

- We are testing artificial discussions on each page in the form of comments on each page. We created over 1415 comments (and counting), all with unique names, personas, point of views on the subject. These comments might contribute helpful material. While the debate is still up on if comments are the result of recurring visitors or if they are specifically being rewarded, we wanted to test them out. If the Helpful Content AI classifier is rewarding helpful information within comments, these should be beneficial.

Update: *While Google hasn't had time to fully process every quote this has not yet had a significant impact on rankings. I expect to get conclusive results shortly.*

Test #4 (On-Going)

- We are testing altering the ratio of SEO focused articles (targeting a specific keyword) versus articles that do not target a keyword. To accomplish this, we added nearly 100 articles of relevant, non-keyword focused articles to a website. The theory is that perhaps Google is targeting sites that only write about keywords is being tested here.

Test #5 (Preparation)

- We will be testing narrowing the scope of content. Reducing the width of certain topics and focusing more on a specific topic as determined by the title of the page. While this is a fairly large endeavor, it will be worth it to test the performance of different types of content. This would entail removing large portions of each page, distilling it down to only the most essential content while preserving a high entity density.

Test #6 (Optional)

- We plan on testing altering the ratio of helpful pages to non-helpful pages by potentially introducing a discussion forum and seeding it with AI content. Alternatively, we might consider flooding a site with narrowly focused content designed to appear authoritative. This could potentially shift the ratio of helpful to non-helpful pages according to the AI classifier.

Test #7 (Optional)

- While it is not recommended, we might test sending fake navigational queries and visitor sessions over an extended period of time visitor signals to a site. This would, in theory, make us pass the helpful content check by virtue of simulating recurring visitors. A branding campaign to increase the quantity of site searches might be easier.

10. How To Recover From The Google Helpful Content Update

(Tentative Step by Step Instructions)

Until I repeatedly recover websites impacted by the Helpful Content Update, **I can only tell you what I would personally do if I had recover my own website.** This will aim to cover all 3 theories.

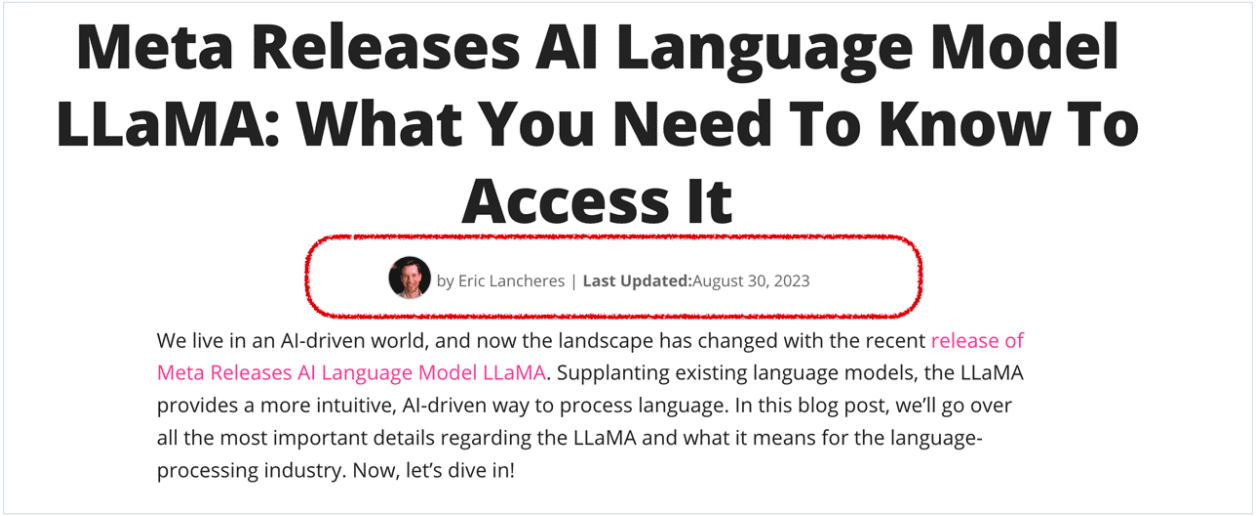
Proceed at your own risk. I have no control over your site or Google, you are 100% responsible for any changes you make to your site.

Here's what I would personally do:

1. I would start by listing the essential trust elements on each articles, specifically:

- Author Name
- "Reviewed By" Name (Optional)
- Last Update Date / Publish Date

These are the bare essentials that I believe should be listed alongside any article and unfortunately some sites do not do this.



2. Clickable "Author Name"

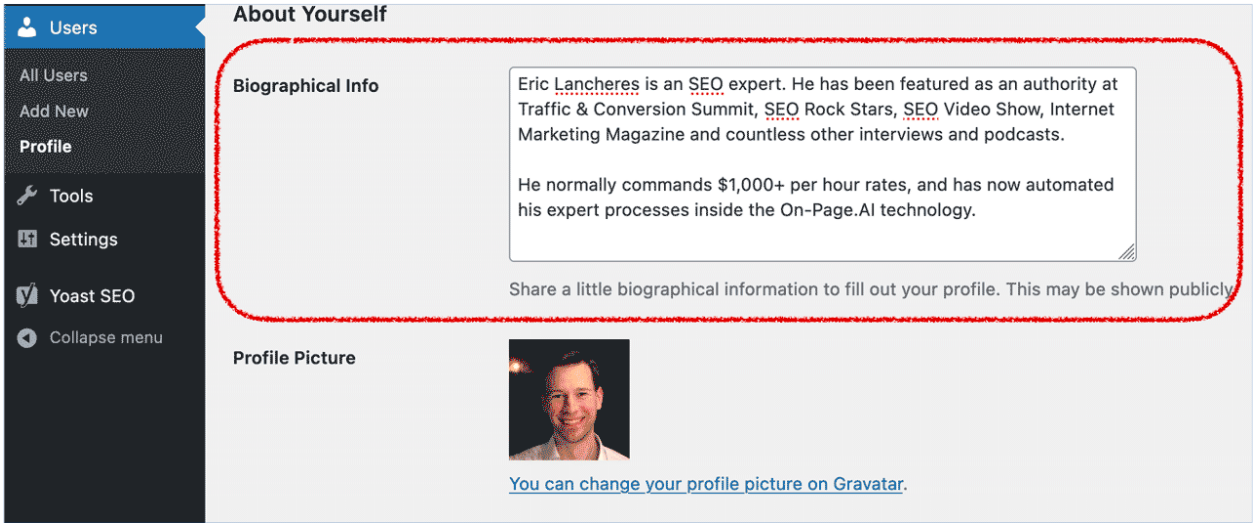
I would make the author name clickable and upon following the link, we would find a detailed author bio.

3. Create a carefully crafted author bio that demonstrate expertise

Within WordPress, you can create a bio for a user and include as much content as you want. To create my bio, I would first ask myself the question: "*Which type of person is qualified to write about this topic?*"

For example, if I have a home website, the ideal author might be a contractor. If I have a dating site, then a relationship coach might be an ideal candidate. Then I would enlist the expertise of a real expert for my site. (*Keeping it legitimate, I'm not advocating to make up titles.*)

Then I would create a bio clearly outlining the expertise and experience of the author. I would include certifications, qualifications, years of experience, specific fields of study within the bio. If you find it challenging to articulate your own accomplishments, consider seeking assistance from a trusted colleague or friend.



(My sales & copy writer wrote this bio for me)

4. I'd prominently display an "Author box" alongside every article.

5. For every article, I would ask myself: "What is the main question that the reader wants answered when they land on this page?" Then I would create a quote by an authoritative body providing an answer to this question, listing the answer near the top of the page. This way, the AI classifier should always find "helpful and trustworthy information".

For example, if I have a page on cleaning wood floors, I might have a quote such as: "*According to the National Woodflood Association Of America, cleaning wood floors once a week...*" near the top. This would add another answer to the page from a different perspective.

In the event that Google is searching for a ratio of narrowly focused content, this would also serve to increase the quantity of content focused on the answer.

6. I would set up my website to be a place of discussion and community.

I would open up comments and seed hundreds of friendly comments after every article to encourage on-going discussion. Even if this isn't a direct signal, preexisting comments encourages users to leave additional comments. This can encourage users to return to the site.

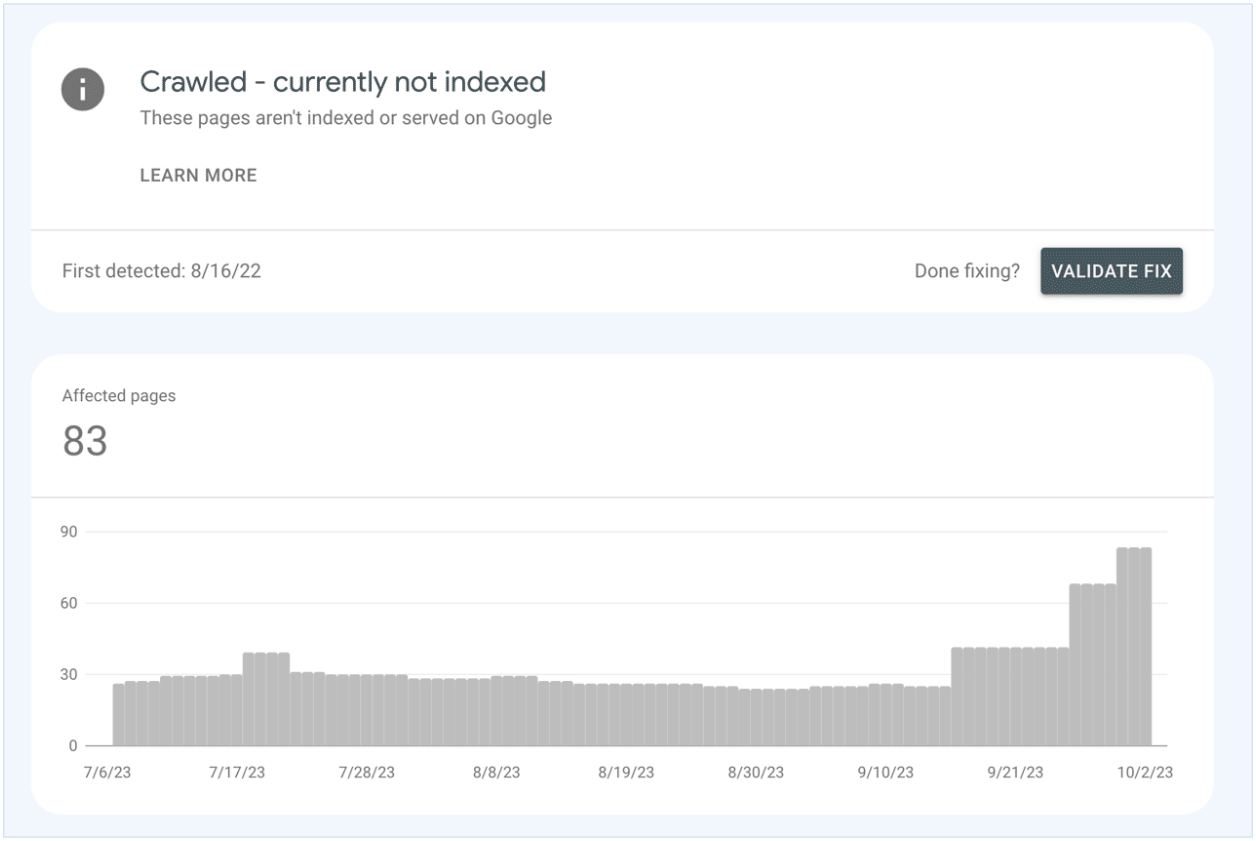
7. I would review the quantity of ads that I have listed on each page.

Checking both desktop and mobile versions. If it's an excessive quantity, I would reduce the quantity of ads per page by a significant amount.

8. I would go into Google's Search Console, identify the pages that are listed as "Crawled but not indexed" and try to identify any pages that might be considered to be thin or duplicate content.

I would consider improving or outright removing thin content as it can hamper a site's performance.

While on the topic, I would also prune empty (or very small) category pages, tag pages. I would remove date archives and I would disable search pages from being indexed.



Then I would wait at least 3 weeks for Google to re-crawl my pages and recalculate my Helpful Content Score.

9. Following that, I would return to my content and begin pruning large pages to be more focused on the specific topic outlined in the title tag.

For example, if I'm writing about "eHarmony Review", I would omit discussing the "history of EHarmony" and instead, focus exclusively on the "review" aspect of the article.

Trimming down articles can often lead to removing necessary entities (relevant words) so **I would constantly check the entity density of the article to ensure I'm maintaining a higher than average entity density** even as I remove superfluous portions of the article. Entity density can be maintained while trimming articles and I believe it is still immensely important.

10. I would start adding topically aligned content that does not focus explicitly on high CPC keywords.

Instead of writing an article on "*how to make money*", I might create an article discussing my journey as an entrepreneur, "*What I Learned Over 20 Years In The SEO Business*". I would aim to create enough of this content to alter the ratio of SEO-focused articles versus normal topical articles. After all, it is difficult for the AI classifier to say that the story about my personal experience in business is unhelpful. I still want to remain on topic to build my topical authority without explicitly going after 'money keywords'.

11. I would set my site name following Google's official documentation. Afterwards, I would begin a campaign to increase the quantity of navigational searches on Google for my website.

```
<script type="application/ld+json">
  {
    "@context": "https://schema.org",
    "@type": "WebSite",
    "name": "Burnt Toast",
    "alternateName": ["BT", "B-T", "Burnt Toast Shop", "example.com"],
    "url": "https://www.example.com/"
  }
</script>
```

Navigational query examples:

- *how to tie a tie Yoursitename*
- *Yoursitename*
- *Yoursitename.com*
- *Yoursitename how to tie your shoelaces*
- *alternate Yoursitename*

I would focus on sustainable, long term methods to increase the quantity of queries for my site's name on Google. (Branding campaigns, forums/community, advertisement, etc)

12. And of course, I would check back here often to see the updated test results and read the most up-to-date helpful content information.

11. Final Words

(References, Further Readings, Support)

This project represents a significant investment—over 180 man-hours from my team and me, thousands of dollars in testing, and numerous late-night sessions.

The reason I have chosen to make this publicly available is to help as many people as possible that were unfairly targeted by the helpful content update. While I understand that some websites deserved the consequences, I also believe that some websites and webmasters were unjustly affected.

To those people, I hope that by sharing some of my work helps you.

If you run an SEO agency, I hope that some of my theories, data and tests help you better serve your clients. You can be the hero that rescues a struggling website.

If you're an SEO professional, then the data might make more informed decisions, saving you time, money and effort.

Should you feel this resources valuable, feel free to share this page within your professional network, within Facebook groups, Slack/Skype/Discord chats and forums.

And should you wish to further support me in any way, I invite you to check out On-Page.ai. Created with Google's best practices at its foundation, the platform aims to optimize your rankings by staying ahead of Google's frequent updates. One of my main motivations for keeping up with the latest SEO trends is to continually refine On-Page.ai, ensuring it embodies the most effective ranking strategies. I'm confident that it can be a valuable asset in ranking your website online.

Here are some useful links:

- Official Google Search Helpful Content Page

- Official Google "How To Create Helpful Content" Questions

- Google "Debugging Search Traffic Drops"

- Google Trial Documents (ie: Search Is Magical)

- Site Names In Google

- Ahrefs Tool

- On-Page.ai

Should you wish to contribute any discoveries, case studies or any corrections, please email me at team@on-page.ai

Sincerely,

- Eric Lancheres