5 ways to use the Wayback Machine for SEO

Sometimes a simple tool can give you incredibly powerful insights.

The Wayback Machine is one such tool.

The Wayback Machine takes historical screenshots of web pages and stores them in its public database. Anyone can use the Wayback Machine to view previous versions of pages or entire sites.

Here are five smart ways you can use the Wayback Machine for SEO.

Get the daily newsletter search marketers rely on.

---

1. Find legacy URLs from old versions of the site

One of the most useful ways to use the Wayback Machine is to find historical URLs that have never been redirected.

The Wayback Machine collects information about your site throughout time. So it might have access to URL data from 10+ years ago.

This is especially important for sites that have been around for a long time. It’s likely that the stakeholder who managed the site years ago has changed the company or left roles and may not have used SEO best practices during site migrations.

The Wayback Machine can be a lifesaver here. You can quickly find old URLs that were never redirected to live versions.

For example, the "Headphones" page from Bose (http://www.bose.com/products/headphones/) from 2003 was never redirected:

Using the Wayback Machine, it's easy to discover legacy versions of key content from previous site versions. You can then find URLs to redirect that you would have likely never discovered otherwise.

Want to take this to the next level? Read Patrick Stox's article on using the Wayback Machine API to find historical redirects. By querying the API, you can bulk export legacy URLs. This can be much more efficient for larger sites.

2. Find previous page content

Website content changes over time. This happens for a variety of reasons (e.g., SEO, CRO, site migration or highlighting different aspects of a product). There is always an inherent risk of any content changes, especially if they're significant.

This is where the Wayback Machine comes into play.

If you're seeing large losses in rankings after changing content, you can check the Wayback Machine to surface previous versions of old pages. Restoring the content to its original version could help your content regain lost visibility.

For example, NYMag's "Best Pillows For Neck Pain" article has lost organic visibility since mid-2020 for terms such as "neck support pillow." This has resulted in organic traffic loss over time.

Comparing the page to early 2020, we see that they have changed the content since then. The 2020 version included a quote from a chiropractor from the American Chiropractic Association in the introduction and kept the products above the fold.

However, in the current version, they have added more content to the introduction, pushed the products below the fold and moved the American Chiropractic Association quote further down the page.

While this might not be the sole cause of the ranking drops, looking back at the previous content during peak rankings could help them test restoring some of the content to older versions to see if this helps improve visibility.

3. Finding old robots.txt file

Another great use of the Wayback Machine is checking how your robots.txt has changed from previous versions. This can be particularly helpful during a site migration if your robots.txt file has changed and you don’t have a version of the original file.

Fortunately, the Wayback Machine crawls robots.txt files a lot. Just look at how many times IBM’s robots.txt file was crawled in 2012:

Using this, you can analyze how the robots.txt has changed over time. For instance, IBM's robots.txt file looks quite different from what it used to. Here is the file back in 2012:

Looking at the site's current robots.txt file, you can see that the commands have changed:

Using The Wayback Machine can be an extremely effective way to lookup old versions of your robots.txt file. This is especially useful if the information has been lost during a site migration.

4. What sections competitors are adding to their pages

Sites in competitive spaces routinely add or update content. For your highest priority keywords, your competitors are likely making frequent updates to their pages to try to improve their visibility. It can be difficult to track these changes.

Fortunately, the Wayback Machine allows you to understand what types of updates competitors are making to their content.

For example, we can use The Wayback Machine to look at the best cast iron skillet page from Serious Eats on June 27, 2021:

Looking at the page today, we can immediately see that they

have made some dramatic changes to the page:

By reviewing the existing page, we can see that they have:

- Added an “Editor’s Note” to the top of the article
- Moved the “The Winners” section to the top of the page
- Implemented internal links at the top of the first paragraph
- Made “The Winners” section more visual
- Added an FAQs section

This is extremely valuable information to have when performing a competitive analysis. These changes can now inform the editorial strategy we apply to our own page.

Determining the content differences can be difficult. It requires a manual review. However, you can use tools like Diffchecker to easily spot the content changes.

5. How frequently competitors are updating content

Use the Wayback Machine to determine how frequently competitors update content.

This is especially useful if you're in a SERP landscape where content freshness matters for visibility.

For example, CNET's high-ranking page for The Best Android Phone Of 2022. At the top of the article, you can see the timestamp for when the article was last updated:

Because technology is extremely fast-moving, it's likely that freshness matters for terms such as "best android phones" since the products are frequently changing. Therefore, we might want to research how often we need to update our own content to stay competitive.

Using the Wayback Machine, we can construct a timeline of how frequently CNET updates these articles. By looking at the previous timestamp on the page, we can look up the most recent historical version of the Wayback Machine that precedes that date. For example, to find the update that happened before March 5, 2022, we can look up what the version on March 2, 2022, looked like.

By repeating this process, we can develop a timeline for how frequently CNET updates this page:

Based on the data, it's safe to say that CNET updates this article on a monthly cadence. We might want to apply the same update frequency to our content to stay competitive with CNET.

Back to the Wayback

In a world where the web is ever-changing, the Wayback Machine is invaluable.

You can use this tool in multiple ways to recover lost information and glean insights into the direction of competitor strategies.

Make sure the Wayback Machine is in your SEO toolkit.

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

New on Search Engine Land

About The Author

Chris Long is the VP of marketing at Go Fish Digital. Chris

works with unique problems and advanced search situations to help his clients improve organic traffic through a deep understanding of Google’s algorithm and Web technology. Chris is a contributor for Moz, Search Engine Land and The Next Web. He is also a speaker at industry conferences such as SMX East and the State Of Search. You can connect with him on Twitter and LinkedIn.

https://searchengineland.com/5-ways-to-use-the-wayback-machine-for-seo-383754