MarketMuse & Sitebulb

The Complete SEO Content and Technical Audit Guide

In this guide, we walk through a four-step process for conducting an SEO content and technical audit, and then acting on that insight. Having crawled over 100,000+ sites, who better to help put together this manual than Patrick Hathaway, Director of Sitebulb. If you missed it, Patrick and MarketMuse co-founder Jeff Coyle got into a deep discussion on content audits in a recent webinar.

Basically, a content audit focuses on evaluating the quality, relevance, and effectiveness of a website's content. The goal is to identify issues like content gaps, poorly-performing pages, outdated, irrelevant, or low-quality content, and opportunities for optimization. The process involves collecting data about all the web pages across a website and making decisions about whether to keep, update, merge, or remove pages.

What is a content audit?

Content audits are the convergence of technical SEO and content strategy, where the result is a comprehensive content inventory and corresponding action plan for every page.

A content audit focuses primarily on evaluating the quality, relevance, and effectiveness of a website's content. The goal is to identify areas for improvement — issues like content gaps, poorly-performing pages, outdated, irrelevant, or low-quality content, and opportunities for optimization.

The process involves collecting quantitative data about all the web pages across a website, combining that with qualitative insights, then making decisions about what to do with each page:

- **Keep** – keep the page as it is, no changes required
- **Improve** – work on changes to improve the page
- **Consolidate** – combine multiple pages together
- **Remove** – remove the URL from the site as the content has no value

Additionally, a content audit should also identify gaps in the content coverage, to help inform the content strategy moving forwards.

Step 1 - Understand your content strategy

Your content marketing efforts are underpinned by your strategy, namely:

- Which buying personas you are trying to reach, and the relative important of each to your content strategy
- What messages you are trying to communicate

It is important you have a clear understanding of this before you start any content audit project, as this will play a key role in helping you determine the purpose of each content piece.

You will need to be in a position to evaluate the content and determine

- Which buyer persona each URL is targeting
- Which buyer stage each URL is targeting: Awareness, Consideration or Decision

It also helps to have a good working knowledge of the website you are working with, as it is often unnecessary to audit every single page on a website, and in some cases you will be able to exclude entire subfolders and save yourself unnecessary work further down the line.

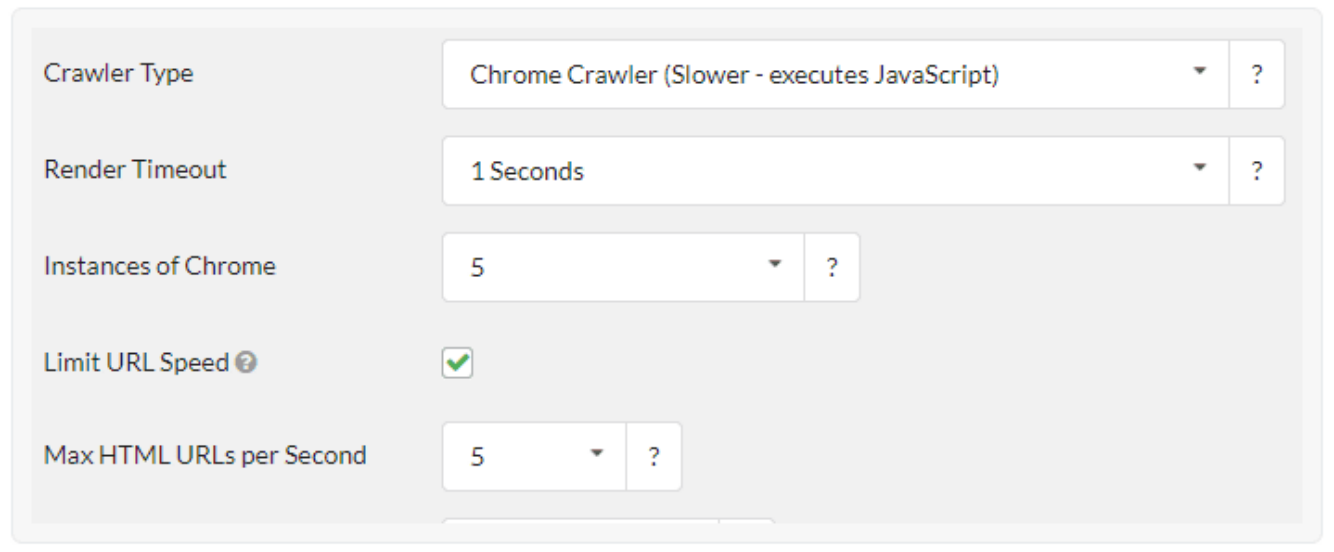
For example, the website could have a /legal/ section that is a necessary evil(!), but that is not designed to be 'consumed' by potential customers in the same way as general content is.

Step 2 - Creating your content inventory

Your content inventory is a big list of every single URL you plan to audit, and to do this stage properly you need an SEO crawler, like Sitebulb.

Sitebulb is one of the leading SEO website crawlers, and as a result of their partnership with MarketMuse, you can get a free extended trial from signing up via this page.

One of the benefits of a tool like Sitebulb is that it easily allows you to crawl and execute JavaScript – which you can do by selecting the 'Chrome Crawler' – this renders the page HTML and reflects what your customers actually see.



As we mentioned earlier, you don't always need to audit every single page on the website. So if you are unfamiliar with the website you are working on, it makes sense to first perform an exploratory crawl for the purpose of URL discovery.

How does a crawler work?

Website crawling works by (programmatically) downloading the HTML content of a webpage, and extracting the internal link references from the HTML until no new links are found.

A simplified version of the process looks like this:

1. The crawler visits the start URL (typically the homepage)
2. The HTML content from the homepage is downloaded, and <a href> links are extracted
3. The URLs from each extracted link is added to the crawl queue
4. The crawler selects a URL from the crawl queue
5. The HTML content from this URL is downloaded, and <a href> links are extracted
6. Any new URLs from the extracted links are added to the crawl queue

Steps 4-6 are repeated until no new URLs are discovered, at which point the full process of URL discovery has been completed. The crawler will continue working through the rest of the URLs to download the HTML content.

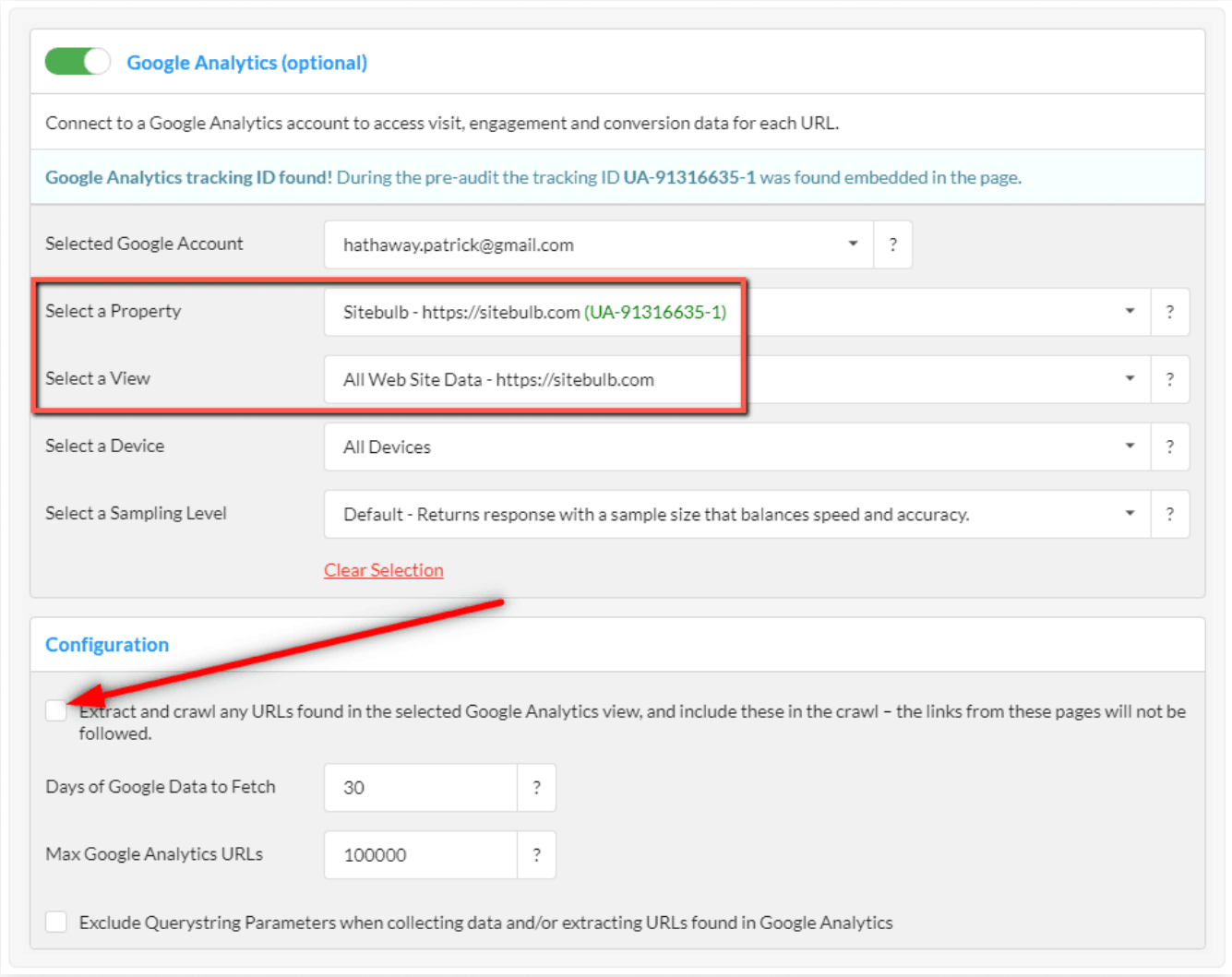
In addition to internal links, the crawler is also extracting key information from the HTML of each page, which it will then use to form the website crawl data.

Performing a complete crawl

When performing a complete website crawl, you don't want to limit the crawler, you want to enable it to find as much as possible (the only exception to this is if you know the website well and know you wish to exclude certain sections). Then, you can look through the data afterwards and figure out which bits you don't really need.

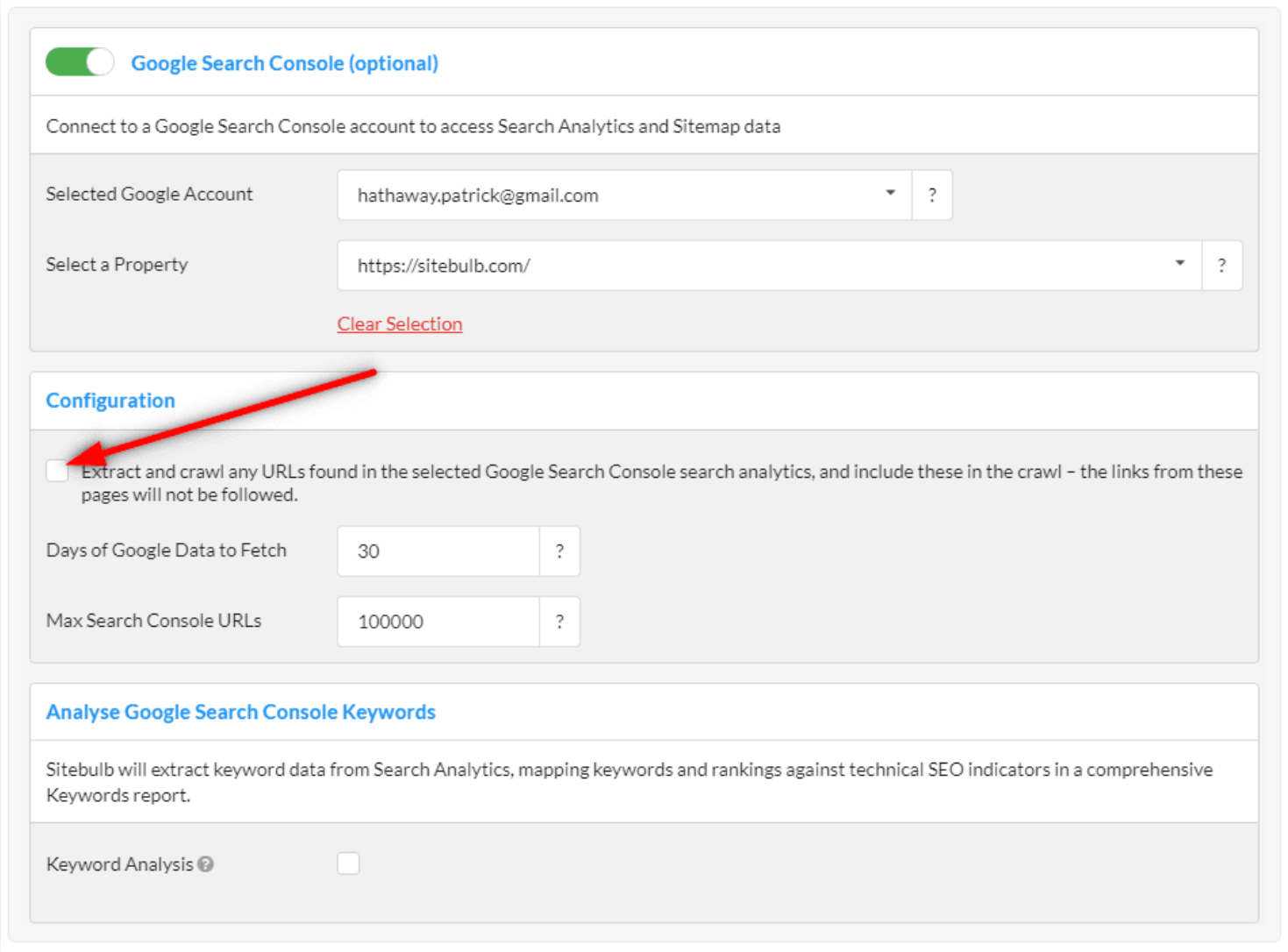
This means you give the crawler all possible crawl sources, including:

Google Analytics
Adding Google Analytics data means that when the audit is complete, we have accurate data available regarding organic traffic and key engagement metrics.



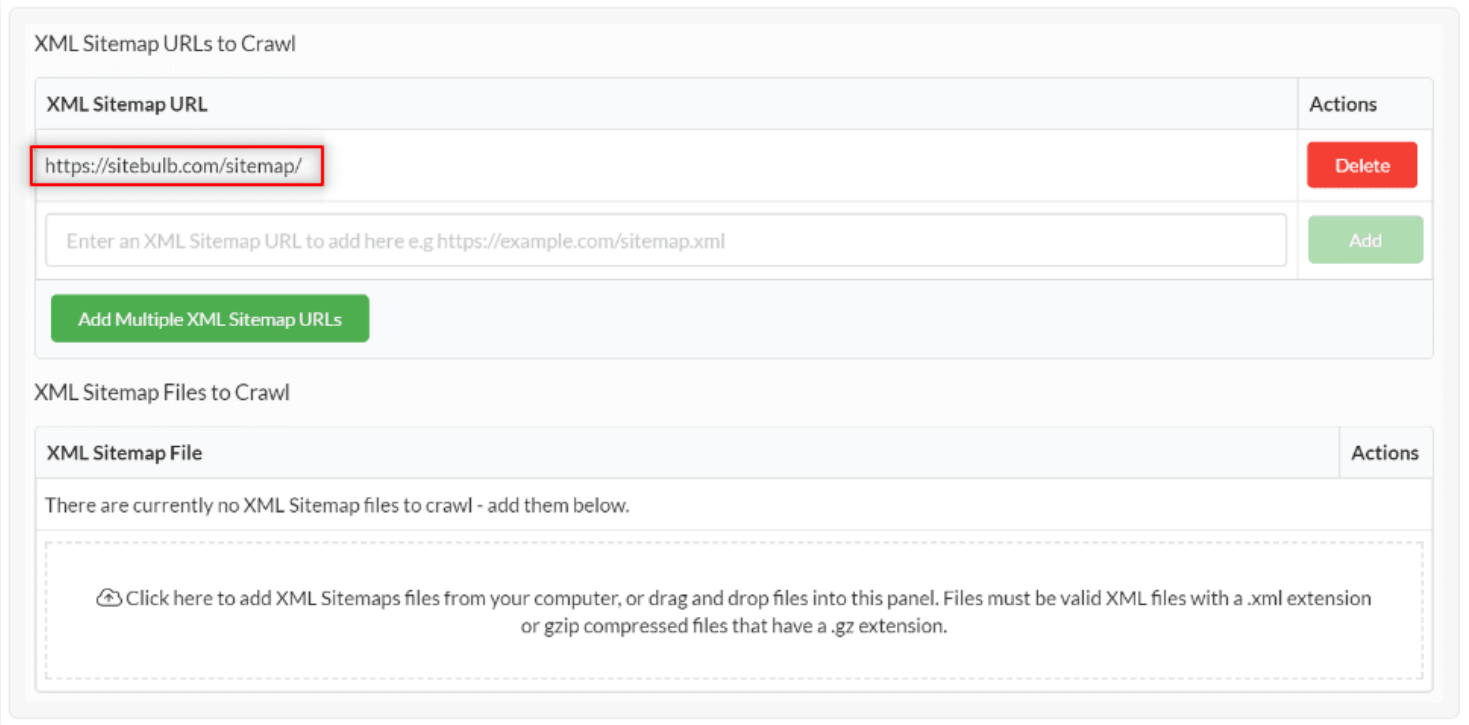
Google Search Console

Adding Google Search Console data means that when the audit is complete, we have accurate data available regarding organic rankings, clicks and impressions.

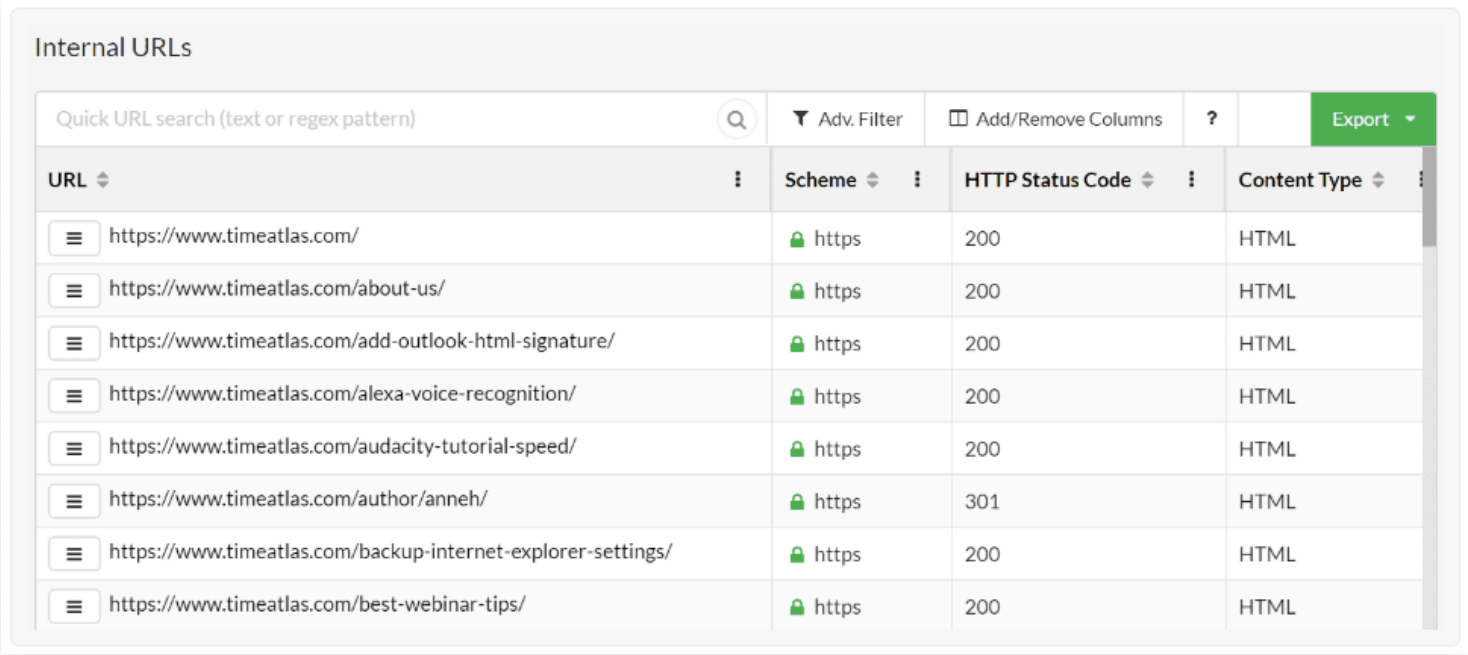


XML Sitemaps

Add in all XML Sitemaps that exist for the website.



The point here is to try and include the entire universe of URLs that exist for the website, so that you end up with a big long list of all the URLs on the site.

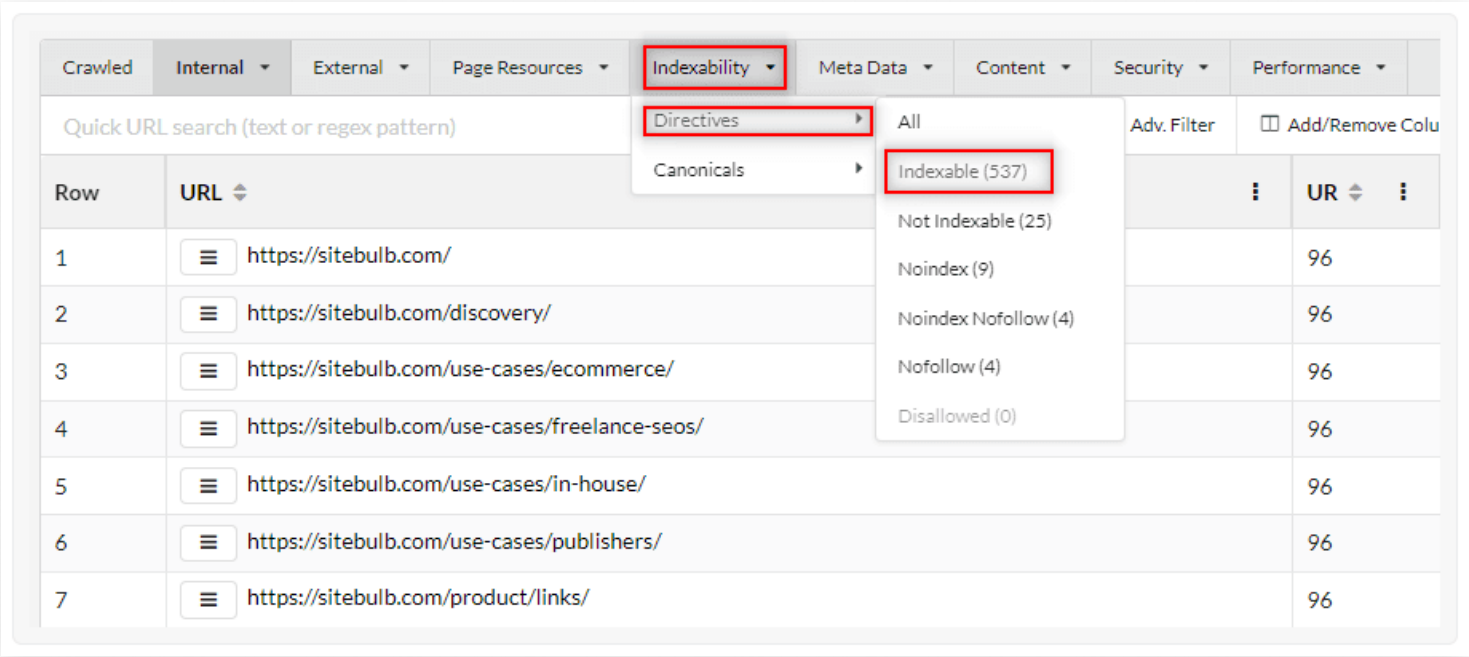


Filtering out unwanted pages

For the purpose of a content audit, you are typically only interested in pages that are indexable, which means that search engines are able to index and potentially include the pages in their search results.

Indexable pages are:

- URLs that return a 200 status code
- URLs that do not have a canonical pointing at a different URL
- URLs that do not include a noindex robots directive
- URLs that are not disallowed in robots.txt



At this point, you will now have a filtered list of all the URLs you wish to consider, along with important quantitative data:

- Title, h1 and meta description
- Word count of the content
- Crawl depth of the URL
- No. of internal incoming links (+ URL Rank)
- Visit and engagement data (from GA)
- Ranking, clicks and impressions data (from GSC)

At this point you will want to export this data into spreadsheet format, to add in the extra qualitative layers. You may also wish to supplement the data with additional 3rd party metrics, such as Majestic or Ahrefs backlink data.

The extra columns you'll need to add could include:

- Content type
- Buyer persona
- Buyer stage

Populating this data will take a reasonable amount of manual inspection, although it won't take long before you start to spot patterns in the URL paths that can help you make decisions in bulk.

	A	B	C	D	E	F
1	URL	Content Type	Buyer Persona	Buyer Stage	UR	Crawl Depth
2	https://sitebulb.com/	Homepage	N/A	Awareness	96	0
3	https://sitebulb.com/discovery/	Commercial	Primary	Awareness	96	1
4	https://sitebulb.com/use-cases/ecommerce/	Commercial	Primary	Awareness	96	1
5	https://sitebulb.com/use-cases/freelance-seos/	Commercial	Primary	Awareness	96	1
6	https://sitebulb.com/use-cases/in-house/	Commercial	Primary	Awareness	96	1
7	https://sitebulb.com/use-cases/publishers/	Commercial	Primary	Awareness	96	1
8	https://sitebulb.com/product/links/	Product	Secondary	Consideration	96	1
9	https://sitebulb.com/features/audit-comparison/	Product	Secondary	Consideration	96	1
10	https://sitebulb.com/use-cases/agencies/	Commercial	Primary	Consideration	96	1

Assessing your data

Finally, you will also want to add in a 'Recommendation' column, that should then be filled in with one of the four status options we defined earlier on.

Using all the data you have collected, make an assessment about each page and assign one of the following statuses:

- Keep
- Improve
- Consolidate
- Remove

It is normally also helpful to have a 'notes/instructions' column alongside, to add more depth and clarity to the recommendation.

This is the most time-consuming step, but also the step with the biggest potential impact. The recommendations made during this stage of the process will help define your page-level content strategy moving forwards.

We'll go through each of the options to help you understand which best fits for each page.

Keep
Select this option when the page is performing to expectations, and no changes are required. This is the case when the page content remains topically relevant and does not need to be updated to bring it up to date.

Common examples: Case studies, evergreen 'pillar' content, FAQs, recent blog posts.

Improve
This is the right option for pages that are not performing well enough, are not correctly targeted or are out-of-date.

Examples of how you can improve pages include:

- Adjust the persona targeting
- Adjust the keyword targeting
- Add or update a CTA
- Add extra content elements (e.g. video)
- Improve internal linking (this is again where Sitebulb can help)
- Re-write sections to bring them up-to-date
- Re-do outdated screenshots or images

We will explore further below how you can use Marketmuse to improve content to make it more competitive in the SERPs.

Common examples: Low traffic pages, older blog posts, pages with poor conversion metrics, good pages with few incoming internal links.

Consolidate

Choose this option when you have multiple pages that are performing the same task or targeting the same set of keywords. If you have any two pages that are designed to satisfy the same keyword intent, then these pages have the potential to cannabalize one another in the search results.

The amount of work required to consolidate two pages will depend on how similar the pages are in the first place. If they are almost identical, then you should pick the stronger page, and add a 301 redirect from the weaker page into the stronger page.

If there is some overlap, but there is still some valuable content on the weaker page(s), then the best course of action is to take this content and add it to the strong page, before removing/redirecting the other page(s).

Common examples: Articles with similar themes, old blog posts, thin content pages.

Remove

This option is for pages that provide no ongoing value, or where the resources required to improve them would outweigh the benefits of doing so.

Candidates for removal can be easily identified by looking at the data you have collected:

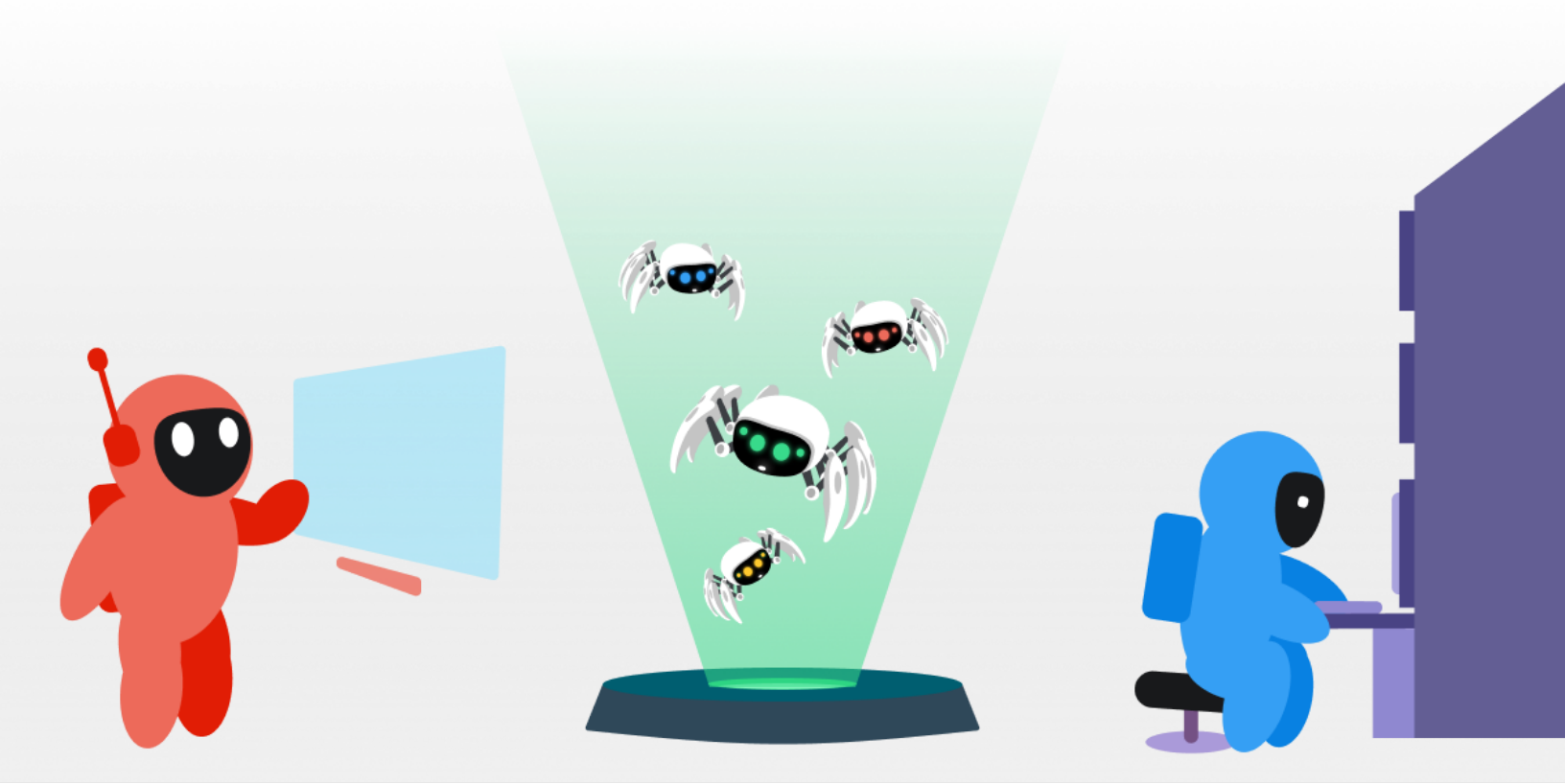
- Pages with very little content (e.g. less than 150 words)
- Pages that did not receive any search traffic
- Pages that didn't receive many impressions
- Pages that rank for irrelevant keywords
- Pages that aren't targeted at any of your buyer personas
- Pages that do not have any incoming backlinks

Whatever you do, don't make decisions based on one dimension alone. A comprehensively written high-quality page can still receive little search traffic, or links, yet still fill a vital role in exhibiting your site's expertise. Removing these pages could have a negative effect on your topical authority.

Removing a page often means deleting it from your CMS, so that it now generates a 404 (Not Found) or 410 (Gone) HTTP response. If this is not carried out carefully, it will then generate 'broken links', as the website will still contain internal links that point at the deleted page – which causes a bad experience for users and search engines alike.

To ensure that this does not cause ongoing issues, you should remove all link references to the removed pages. If you are not sure how to find these, run another crawl with Sitebulb, and it will find them for you (guide on how to find broken links with Sitebulb).

Common examples: Location 'doorway' pages, discontinued products, duplicate content, old 'event' pages.



Step 4 - Creating a plan of action

Bigger sites offer more opportunities but come with additional challenges. How do you know what to do and what actions are most likely to succeed?

Resource-constrained content teams don't have the luxury of guessing and this is where MarketMuse's predictive personalized metrics come into play. Unlike the generic metrics that most SEOs use, these ones are personalized — based on your site, its content, and performance.

Take that long to-do list of content you want to improve and prioritize it based on Topical Authority and Personalized Difficulty.

Topics | Search for a topic, URL or page title... | Export

Highest Topic Authority — Non-branded | Configure Table | 2 Filters | Show Archived Items | Add to inventory

TOPICS	TOP RANK	TOP RANKING URL	TOPIC AUTHORITY	DIFFICULTY	PERSONALIZED DIFFICULTY	TOPIC VOLUME
topic clusters	1	https://blog.marketmuse.com/what-are-topic-clusters/	47	48 Medium	1 Easy	480
content clusters	8	https://blog.marketmuse.com/what-are-topic-clusters/	43	49 Medium	5 Easy	210
content brief template	1	https://blog.marketmuse.com/5-elements-of-a-highly-successful-content-brief/	43	47 Medium	4 Easy	260
content brief	1	https://blog.marketmuse.com/what-is-a-content-brief/	42	41 Medium	1 Easy	720
seo outreach	1	https://blog.marketmuse.com/link-building-outreach-tips-tools-and-tactics/	37	38 Medium	2 Easy	390
topic cluster	1	https://blog.marketmuse.com/what-are-topic-clusters/	34	56 Medium	22 Easy	390
core web vitals seo	1	https://blog.marketmuse.com/core-web-vitals-overview/	34	63 Medium	29 Easy	210
user intent	3	https://blog.marketmuse.com/what-is-user-intent/	34	56 Medium	22 Easy	590
content strategy	9	https://blog.marketmuse.com/what-is-content-strategy/	34	82 Hard	48 Medium	8,100
link building outreach	4	https://blog.marketmuse.com/link-building-outreach-tips-tools-and-tactics/	34	54 Medium	20 Easy	210
authority sites	18	https://blog.marketmuse.com/how-to-build-an-authority-website-for-the-longterm/	33	66 Medium	33 Easy	210
pdf seo	7	https://blog.marketmuse.com/why-pdfs-are-not-ideal-for-seo/	32	48 Medium	16 Easy	260

Results: 1 - 100 of 780 | Previous | 1 2 3 ... 8 | Next | Rows: 100

Topical Authority is your competitive advantage — you want to do things that are most likely to succeed. Personalized Difficulty indicates the extent of work required. That could be a simple update to the page, but it could mean creating a whole new page and even additional supporting content.

Again, these are personalized metrics and are based on your site and its existing content. That means your plan will be unique to your site.

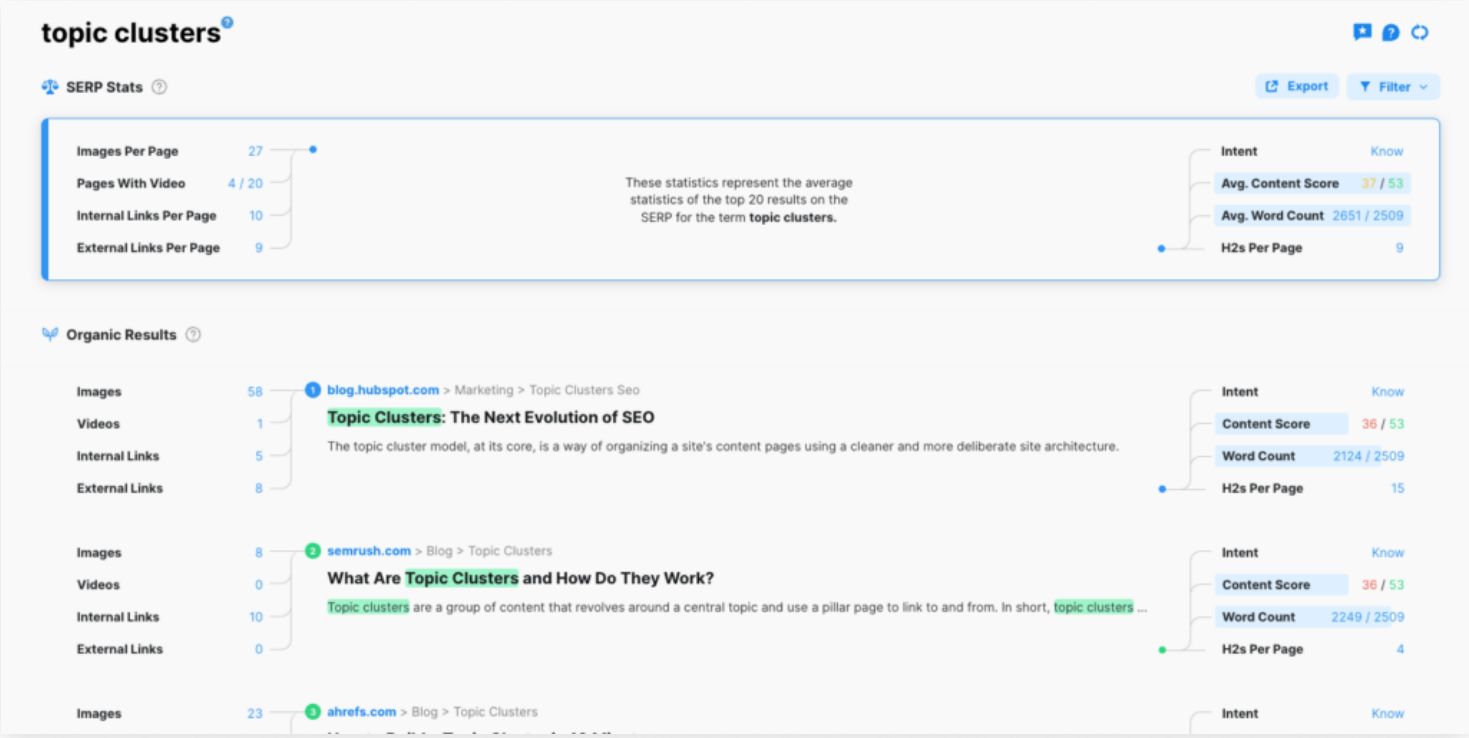
Research And Analyze The Main Topic Of Your Page

A page can rank for many different topics so your first objective is to ensure it's optimized for its main subject. So you'll need to:

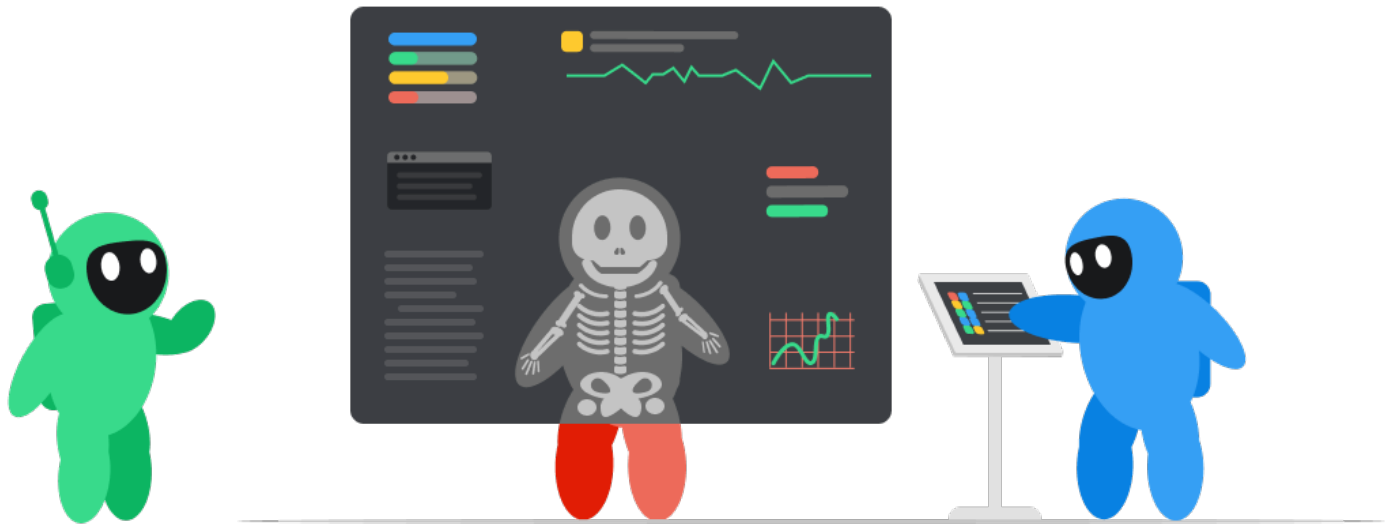
- Determine what the page is about.
- Research that topic.
- Analyze the page.

In researching the topic, use SERP X-Ray to understand how top-performing content is constructed. Some things to look for:

- Primary Intent (informational, transactional comparison, etc.) that Google serves.
- Whether the intent is fractured (multiple intents seen in top 20 results).
- Content Score and Word Count vs. the Targets for these values — help understand the content quality playing field.
- The types of pages (product, service, landing page, blog post, etc.) appearing in the SERP.
- Images and video usage — affects your content budget and may provide a differentiation opportunity.
- Internal linking — may reveal possible content clusters.

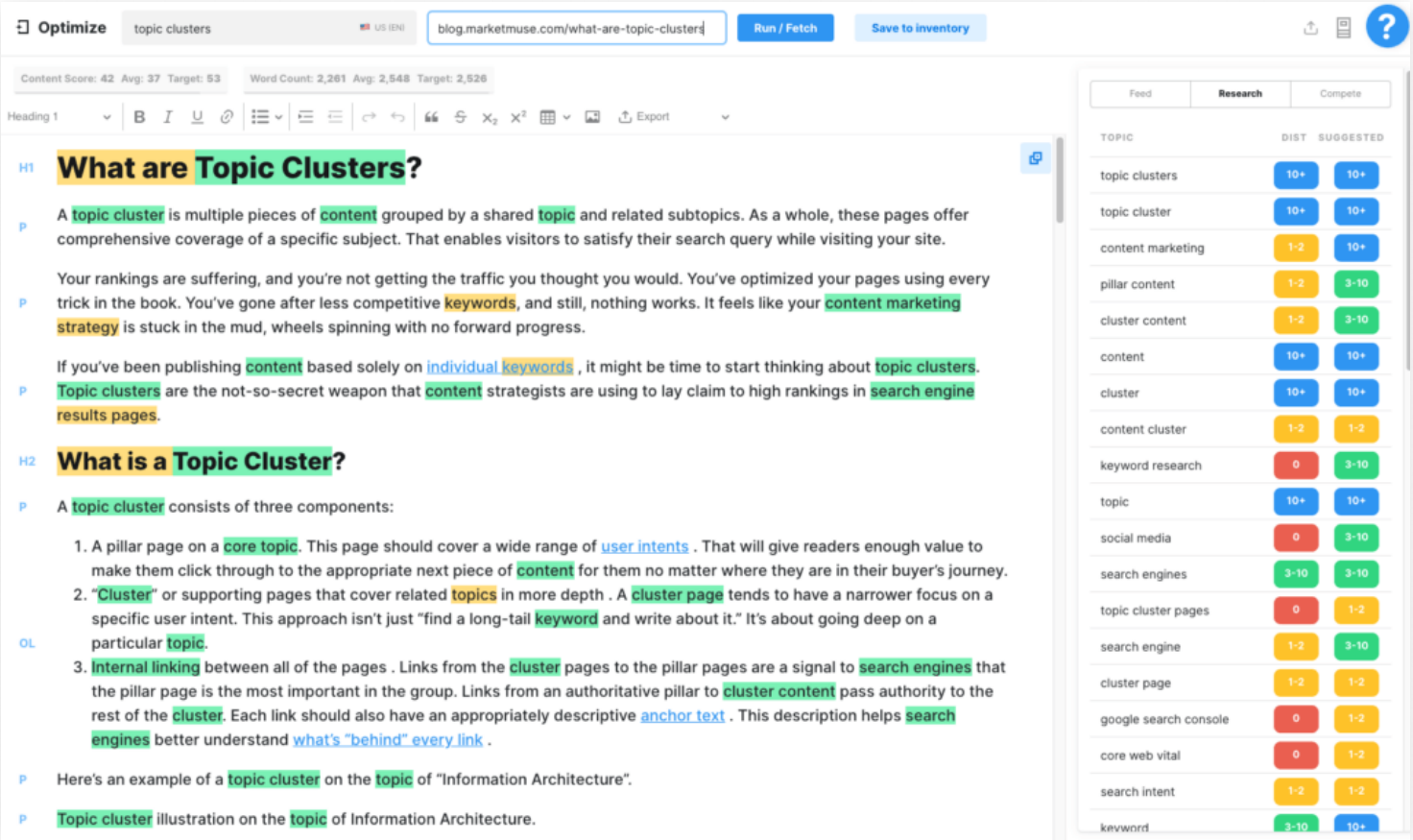


Then use Heatmap to conduct competitive analysis and find ways to differentiate your content. Each of the top 20 results has a content score just like in SERP X-Ray. But here it's easier to conduct a competitive content quality analysis. Pay special attention to the topic gaps as those offer an opportunity to differentiate your content and make it stand out from the crowd.



Optimize the page for its main subject

Use Optimize after you've done the research to better understand the topic and your competition. You want to update the page ensuring it at least meets the Target Content Score and falls within a reasonable range of the Target Word Count.



Content Score is calculated in a way that encourages broad topical coverage. Meeting the Target Content Score, within the Target Word Count will result in a concise yet topically rich piece of content. The suggested ranges for topic mentions serve as a guide to better understand how deeply to cover an issue.

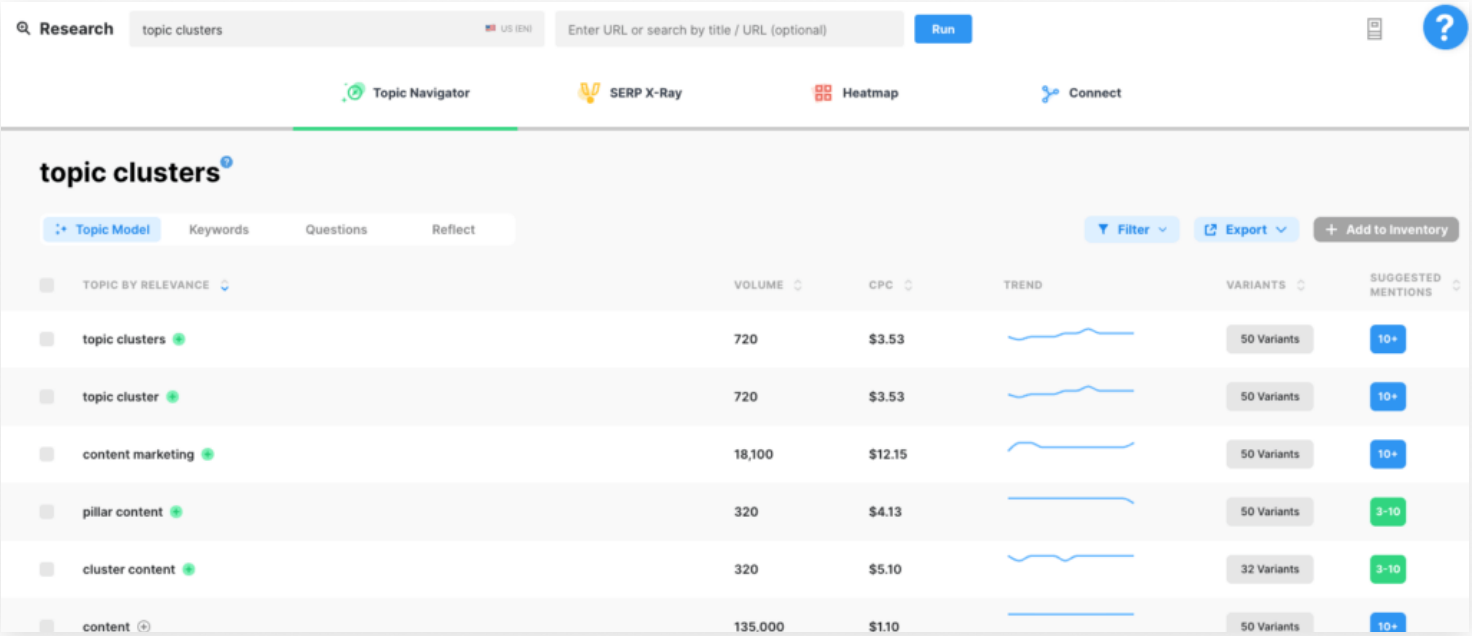
Every situation is unique but here's a basic approach you can use:

- Check that your term usage aligns with the model — a common issue is using acronyms instead of complete spelling. Remember, MarketMuse's topic model is based on hundreds of articles written on the subject. So it's best to follow convention and use words that resonate with your audience.
- Look for opportunities to expand coverage of an existing topic — add content and provide additional clarity. Sometimes a sentence or two is all you need but other times it may require a paragraph.
- Add a section to cover topics you may have missed. Heatmap is especially helpful in surfacing topic gaps, giving you the opportunity to differentiate your content.

Update the page for related ranking topics (and create supporting content)

A well-written and highly authoritative page can rank for many other terms as well. But they may not align well with the topic of the page. If you find such a term that's valuable and want to improve its ranking, don't optimize the page for that term. Optimizing a page for multiple terms that are loosely aligned will end in disaster — you end up changing a great page about one subject into a mediocre page about many, trying to serve competing intents.

Instead, create a new page specifically targeting that phrase (use Research and Optimize).

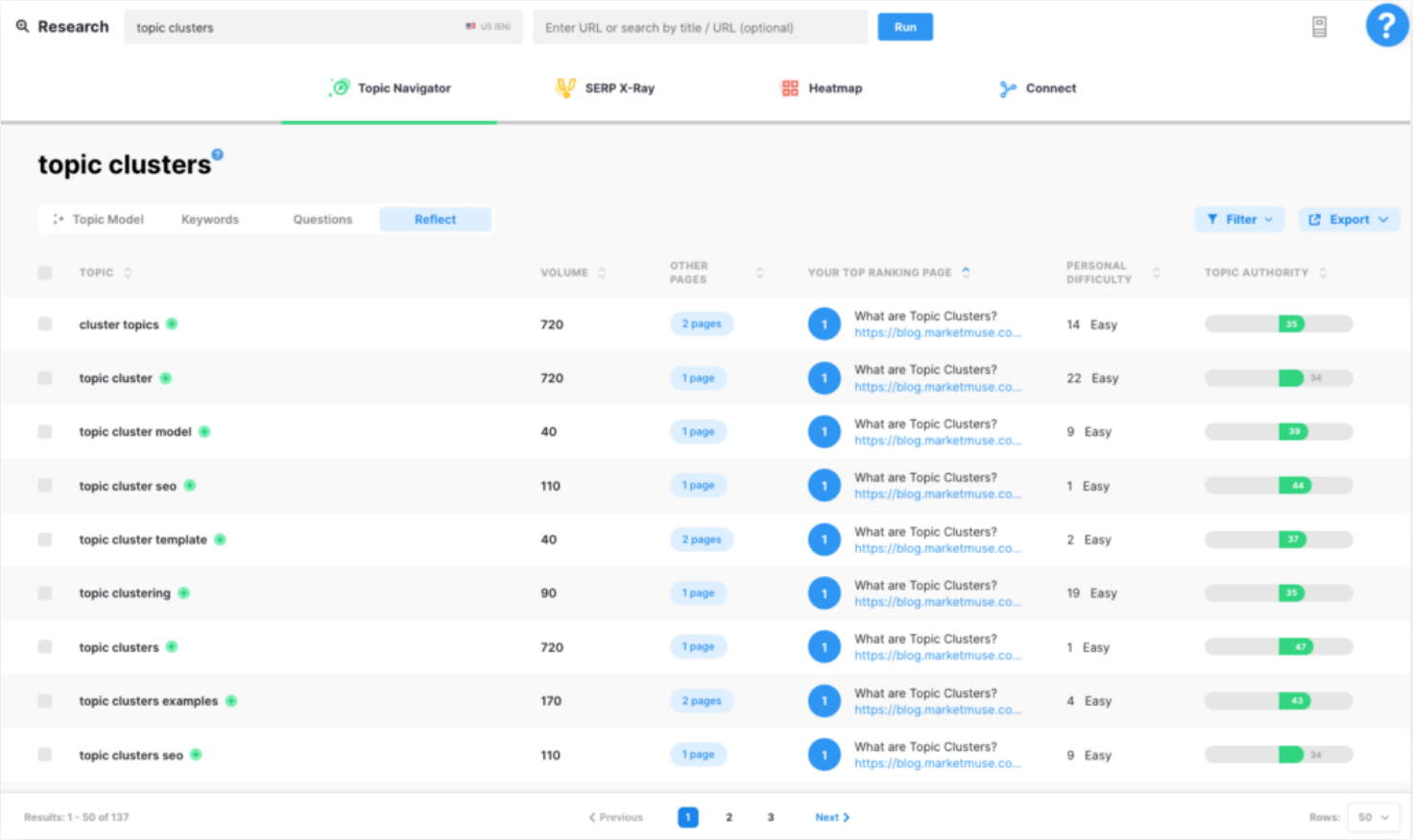


Remember to update the existing page with some connective content. It could be a sentence, a paragraph, or a section — anything to create a contextual relationship. Link to that newly created page to cement that context and pass authority to your new creation.

This assumes you have some Topic Authority (at least 10+) and acceptable Personalized Difficulty (anywhere between 20 and 35).

Create content clusters

When Personalized Difficulty goes north of 35, you'll need to update and expand your current content cluster. In cases like this it's not enough to simply optimize the one ranking page. Reflect shows your current coverage of the topic cluster and updating existing content is always a good place to start.

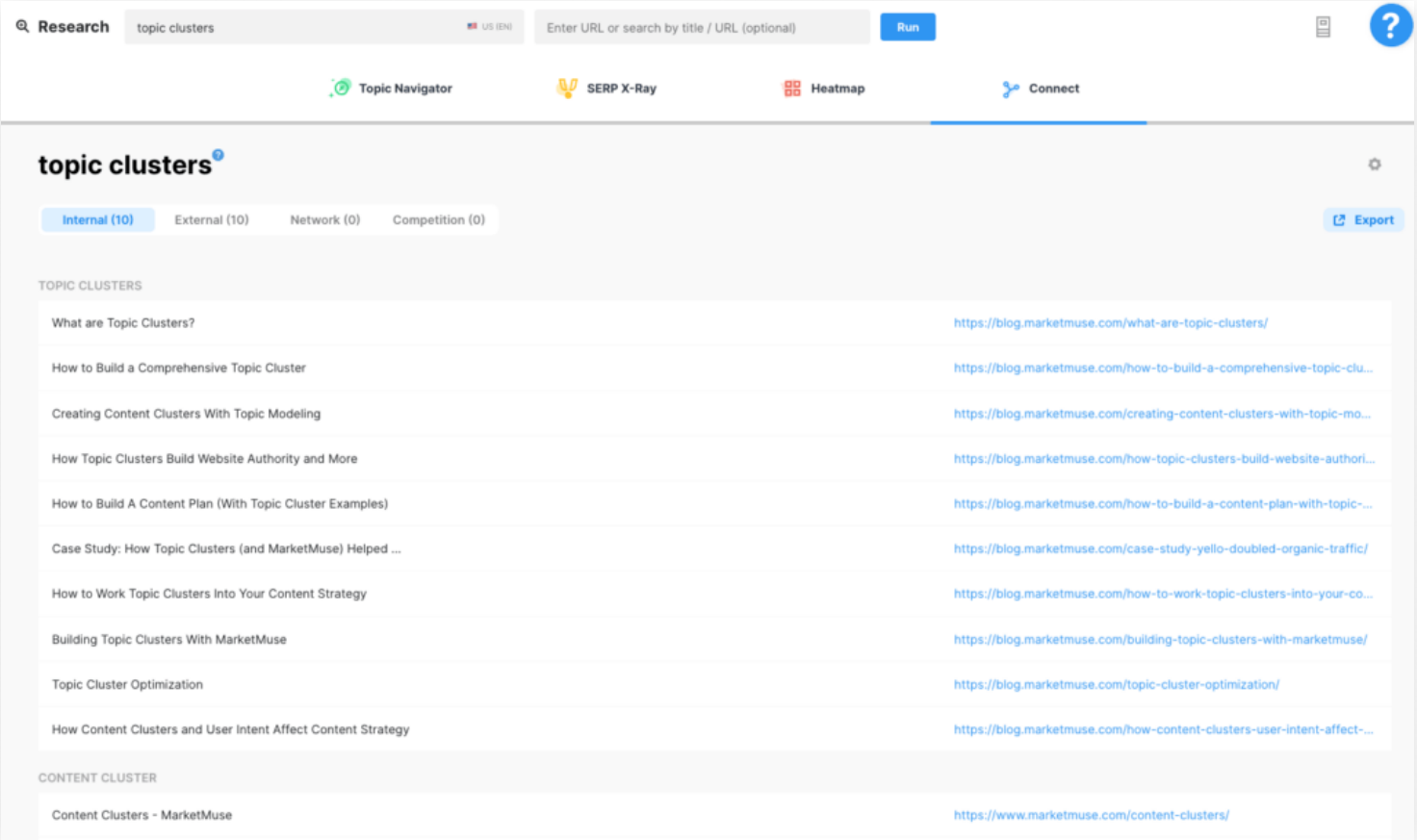


If you have significant coverage, as in hundreds of ranking topics, consider prioritizing using a combination of Personalized Difficulty, Topic Authority, Rank, and Volume.

Looking to add new content pages to the cluster?

You can get some good ideas by reviewing variants of the head (main) term of the cluster. This gives you a better understanding of the buyer journey and what else searchers are looking for.

Take it a step further and make sure you internally link these relevant pages together. Whenever you're creating new content or updating an existing page, use Connect to get linking suggestions. Simple, yet powerful, it ensures a tight connection between semantically related content.



Connect uses the topic model to find appropriate anchor text and analyzes your site to find the best matching pages. The more content you have the more choices are available — MarketMuse provides up to 10 anchor text suggestions each with up to 10 links. This gives you a great amount of editorial discretion while still being optimized for SEO.

Use content briefs

All the information garnered through Research can be bundled together into a set of instructions for a writer — called a content brief. Here are a few reasons to consider using MarketMuse content briefs:

- Eliminates confusion and the back and forth conversation between writer and editor.
- Ensures a writer produces top-quality content on the first draft and not the third rewrite.
- Serves as a source of truth to which everyone can refer.

Takeaways

A content audit is a critical step in ensuring that your website is optimized for search engine visibility and user engagement. By understanding your content strategy, creating a content inventory, and performing a comprehensive audit of each page, you can identify issues and opportunities for improvement. With the right tools, such as Sitebulb and MarketMuse, you can easily complete a content audit and develop a comprehensive action plan for your website.

MarketMuse

Publish better every time.

Let us show you how MarketMuse helps thousands of marketing teams turn content into a growth channel.

Let's talk