Dark Web Search Engine

Dark web search engines exist, but even the best are challenged to keep up with the constantly shifting landscape. The experience is reminiscent of searching the web in the late 1990s. Even one of the best search engines, called Grams, returns results that are repetitive and often irrelevant to the query. Link lists like The Hidden Wiki are another option, but even indices also return a frustrating number of timed-out connections and 404 errors. Many so-called Dark Web search engines are really just repositories of links. This is actually how early search engines on the internet worked. More like a giant phone book than a web crawler that indexed the contents of sites. Then, of course, there are search engines on the Dark Web that search the surface web. In other words, they provide a super-secure way to search for things on the regular internet that you don’t want to be attached to your history or identity. So adjust your expectations a little of what it means for something on the Dark Web to be a search engine and feast your eyes on these excellent Dark Web destinations, in your search for hidden network content.

When we search about something on any search engine, it simply displays up a few results consisting of about 10 links and we found at least one link to satisfy our searched term most of the times. This is call as simple searching or we may web surfing. This way we are simply surfing over the web pages using a traditional search engine. But what is exactly meant by Deep Web Search? To explain this, we are going to take help of illustrative examples. We use the Internet, means the Web to explore, learn and find a lot of things. These things include the information gathering, photos and videos gathering, documents gathering etc.

When one make use of Internet to find anything, there are two type of methods that may be used under our today’s scenario. The very first method is to find the relevant information by searching through the Search Engine like Google and then afterwards surfing the web in simple way. The next method is the Deep Web Search which is not known to most of us. Deep Web Search is to browse the web in an advance way to find a kind of hidden information or any other kind of data which we cannot found by simply browsing the web using the search engines. I may also be said that Deep Web means to explore the hidden Internet.

Most of the people think that using a search engine like Google, they may find some relevant information and hence also got satisfied but actually, they don’t know the Internet is not just limited to it. Sometimes, we come across some websites which are itself a search engine such as the collection database sites.

**What is Hidden Web Search Engines?**

Both Deep web & Dark Web refers to Hidden Web search engine. The terms “deep web “and “dark web “are often interchangeably used-although they are not the same thing. The dark web is technically a small sliver of the deep web, which accounts for 0.01 percent, but the horror stories you hear about the dark web do not actually occur on the deep web. In fact, the majority of the content on the deep web is very similar to the content found on Google, the surface web. And without knowing it, we use it every day.

**What is Deep Web Search Engine?**

Now coming to understand the meaning of Deep Web from the words itself. The word “Deep” clears and exclaims as the web which is deeply hidden in the depths of web. There may be a lot of reasons why search engines do not index these kind of information in the search engine. There might be the factors like the owners of the deep web content may not like to display up their content publically via the search engines and so on more. Anyhow, we are going to learn now about Deep Web Search Engine in Detail.

Deep Web is the data that is not indexed by a standard search engine like Google or Yahoo. The Deep Web refers to all web pages that search engines cannot find, such as user databases, web forums required for registration, webmail pages and pay wall pages. Then there’s the Dark Web or the Dark Net–a special part of the Deep Web hidden. The Deep Web and the Dark Web are the fascinating subjects for the Netizens. However, when you hear the term ‘Deep Web’ or’ Dark Web,’ you usually classify it into one. If yes, you’re mistaken.

There are also different types of Deep Web Search Engines to research or simply search about different type of contents. Some Deep Web Search engines are meant to simply found the deep web textual content and some to find deep web media content. According to some source, the size and the volume of content of the Deep Web is much more than the normal web which browse in general from day to day. But why we are unable to find this kind of web information? The simple answer it is deeply packed, protected or even hacked.

**What is DARK WEB Search Engine?**

Dark Web is where you can operate without tracking, keeping you completely anonymous. The Dark Web is much smaller than the Deep Web and consists of all sorts of websites selling drugs, weapons, and even hiring murderers. These are hidden networks that prevent their presence on the surface web and their URLs are tailored to (.onion). These [website name].onion domains are not indexed by regular search engines, so you can only access Dark Web with special software-called the TOR. TOR is free, and downloadable by anyone. Many of us heard about the Dark Web when the largest underground online marketplace on the Silk Road was dismantled after an investigation by the federal authorities in the United States. But what if you can still dig in your regular browsers the Darknet content without the need for TOR?

**Why isn’t Google’s deep web search available?**

Google does not provide deep web content primarily because this content is not indexed in regular search engines. These search engines therefore do not show results or scroll to a document or file that is not indexed on the worldwide website.

The content is supported by HTML forms. Regular search engines are scrolling and searches are based on interconnected servers. Interconnected servers mean that you interact regularly with the source, but this does not happen when it comes to the dark web. All is behind the veil and is hidden internally in the Tor network, guaranteeing security and privacy. Only 4% of Internet content is accessible to the general public and the remaining 96% is hidden behind the deep web.

Now, the reason that Google does not collect these data or why profound web content is not indexed is not a secret. These companies are primarily illegal or bad for society as a whole. The content can be porn, drugs, arms, military information, hacking tools and so on. Robots Exclusion The robot.txt we normally use is to tell the website which of the files it is supposed to record and register. Now we have a terminology called “exclusion files for robots.”

Web administrators will tweak the setup so that certain pages do not appear for indexing and remain hidden when searching for the scanners. Let’s look at some of the screwdrivers in the internet.

Some of the more reasons why Traditional Search Engines do not index these type of Deep Web Search content are described below, hence we have described some of the properties of the Deep Web Search Engines. These Deep Web resources may contain complex databases which are not easily understood by the Search Engine bots and thus not indexed also. Such Non-Indexed Content may include Contextual Web, Dynamic Content, Limited Access Content, Non-HTML Content, Private Web Content, Software, Archives etc. Coming a little out from the world of Deep Web Search Engine, and reviewing the Deep Web Search. There is no any such thing that a Deep Web Search Engine is necessary to browse the Deep Web but it the only Best option to browse the Deep Web. We can also browse the Deep Web using a lot of other ways or methods. In other words, we may say that Deep Web Search Engines are one of the Best options to choose to search the Web Deeply.

**Top Deep Web Search Engines**

* **Torch -** Torch has one of the largest search engines in the deep web, as they claim to have an index of more than a million hidden page results. It is one of the oldest search engines. Same as notEvil this search engine also having very easy layout, only having one text box for searching one button. Here you only need put your query on search engine box and hit search button, holla result on front of your eyes. Torch is working just like as Google search engine, when you searched any query that you will got good no or result.
* **DuckDuckGo -** This deep web search engine—which, like many other deep web search engines on this list, also lets you search the regular web—has a clean and easy to use interface and doesn’t track your discoveries. The options for topics to search are endless, and you can even customize it to enhance your experience.
* **Onion URL Repository -** The Onion URL Repository has a massive index of over a million page results and does not set limits on what information it holds close.
* **Uncensored Hidden Wiki -** When you search the deep web occasionally, you’ll discover useful places where you need to be careful. The Uncensored Hidden Wiki is very much one of those locations, an uncensored collection of links and articles that, over the site’s history, have included links to information on criminal activities from drugs to child pornography. The site has cleaned up its act considerably, but there are still links to graphic content and possibly illegal sites to be found. If you can look past those elements, however, Uncensored Hidden Wiki is a treasure trove of deep web information. Inside you’ll discover blogs about Tor, links to deep web email services, and even social networks. Just be careful what you search.
* **notEvil -** This search engine is great because users can skip all the clutter and distraction from surfing the web with no ads. It’s clean and mimics the look of Google. The search engines not for profit’ not Evil’ completely survive on contribution and it seems to receive a fair share of support. Highly reliable in search results, this SE has a highly competitive functionality in the TOR network. This search engine provides all type search result by the help of Query or URL, means if you having query and want to find result related to your query then notEvil can help you. notEvil also can search by the help url, and can find your result very quick. One more thing when you will search any query on this search engine, and you not found any good result, then this search engine chat service can help you.
* **ParaZite -** The deep web can seem like a terrifying place, but part of the fun of discovery is opening doors and not knowing what’s behind them. ParaZite is a search engine that gamified the deep web. Beyond its basic, and useful, search features, it also offers up the chance to gamble by taking you to a random site on the deep web. It’s basically the deep equivalent of Google’s “Feeling Lucky” feature—except using it I was taken to a new email client, a black market site, and an essay on why children are jerks. Make sure you’re using a firewall and VPN before you fire up ParaZite.
* **TorLinks -** The directory in TorLinks has a wide range of intellectual results, with the most notable being the literary section.
* Grams - This is the search engine to end all drug-related hunts. Down in the depths, users are able to search effectively for the dark net drugs that are available for purchase on the majority of web page results.
* **Touchgraph -** Touchgraph gets visual with the deep web scavenger hunt. The algorithm it uses is specifically designed to cluster the relationship between your search results to create a visual result—a creative touch to make searching more exciting.
* **Ahmia.fi -** Ahmia.fi is a great search engine for beginners who are new to the deep web. It takes about five seconds to figure out how the search engine works. Once cracked, scouring the deep web becomes a breeze. Ahmia is another top deep web search engine, which deep web user’s likes to find his required results, I also used Ahmia many time for search. Same as other deep web search engine, Ahmia also offering query searching service, means put your query into search text box and press search button and get result. It offers some great services like Ahmia viewer, add .onion links into Ahmia database, I2P Seraching. Ahmia automatically detect bad .onion links and blacklisted into his database, and also maintain his most link visit charts which you can see by the help of *http://msydqstlz2kzerdg.onion/stats/viewer* links.
* **Yippy -** Like Touchgraph, this search engine also collects your searches to make a common result or pattern, but without the visual aspect. Instead, it’s simple like Google. It is a Meta search engine (it gets its results by using other web indexes), I’ve included Yippy here because it has a place with a device entry for a web client, e.g. email, games, videos and so on.
* **Candle Search Engine -** Candle is another alternative for deep web search engine; this is also popular into deep web community, everyday people use this search engine, I also try this search engine many time for deep web searching. Candle search engine does not allow parentheses, Boolean operators; any type quotes into search query, if you put any of these types’ things into search query, then you wouldn’t get required results. Only you can try simple words, For Example, Today I am searching query “deep web links” then I put my query into search box then hit enter, now I am getting some results on my computer screen, but result have only those type sites which have .onion extension domain.
* **Searx Search Engine -** Searx is another alternative search engine for deep web, this search engine also working on free information, if you search here anything then Searx not maintain your log file. This is simple means no one can trace your all activity which you will perform on Searx search engine, here you can search anything which you want. One thing is very good in Searx, This search engine is open source search engine, you can modify Searx source code according to you and also you can participate in search engine functions enhance program. Searx search engine is a great program which collects data from multiple popular search engine like Google, Bing and Yahoo. But here I saw one common thing if any website have high ranking into multiple search engine then searx give top position into search result. Searx search engine is not able to search .onion domain sites. If you try to search anything about .onion sites related then you cannot get required result. Searx search engine offer one great feature which is file related search, you can select any file related option which type results you want to get.
* **Archive.org -** Maybe this was the very first Deep Web Search Engine, I ever found. And also is the top of the most useful websites ever available on the Internet. According to them, they are the non-profit library of millions of this like books, movies, software, music etc. As mentioned on their homepage, they have about 466 billion web pages saved in their database. They have the screenshot of almost every website ever visited on the Internet. You can just image how big this database. Owing to its bigness, how precious and valuable it is. On this search engine, you can know how different websites, which are now touching the heights of the Internet were actually looks like when they were in the starting days. You can see how Google looks when it was started. You can see how Bing looks when it was started. Similarly, you can see almost each and every site’s history webpages. This feature of this ultimate search engine is popularly known as the WayBackMachine, which may took you to the olden times of the Internet.
* **Vlib.org -** This is also a Deep Web Search Engine with a very much vast Database. You can browse through a lot of categories which includes Agriculture, Information and Libraries, Computing and Computer Science, Regional Studies and much more. This is also one of the oldest deep web search engine on the Internet.
* **Wolfram Alpha: Computational Knowledge Engine -** This Deep Web Search Engine one of the Best ever website I have ever found on the Internet which impressed me in the first visit. As mentioned, it is computational knowledge search engine which will show you the results based on the things related to calculation and computational things. On the official homepage, there have given a lot of things that we can do with Wolfram Alpha. Some of them are mentioned below. Mathematics, Step by Step Solutions, Statistical & Data Analysis, History, Media, Finance, Physics, Engineering, Geography, Chemistry, Education, Web and Computers and much more. Here is an example search made in the Mathematics section of the Wolfram Alpha Deep Web Search Engine. Mathematics, Step by Step Solutions, Statistical & Data Analysis, History, Media, Finance, Physics, Engineering, Geography, Chemistry, Education, Web and Computers and much more. Here is an example search made in the Mathematics section of the Wolfram Alpha Deep Web Search Engine.
* **Deep Web Tech (DeepWebTech.com) -** This is one another example of the Deep Web Search Engine. In this deep web search engine, you may browse multiple other deep web search engines such as the science.gov. Science.gov is also one of the most popular deep web search engine where you can find the information and resources related to US federal science.
* **Kilos -** Kilos is a dark web search engine by all definitions. It was created with the sole purpose of letting users search for “Drugs” on the dark web. It fetches results from all the indexed Darknet markets on the Dark web (currently 6) as well as from forums. Even displays “featured listings” on the left-sidebar.

**How Dark Web Search Engine Works and their way of Indexing**

Search engines like Google are incredibly powerful, but they can't crawl and index the vast amount of data that is not hyperlinked or accessed via public DNS services. However, there are Deep Web Search Engines that crawl over the TOR network and bring the same result to your regular browser. These Deep Web search engines talks to the onion service via Tor and relays, resolve the .onion links and then deliver the final output to your regular browser on the ordinary World Wide Web. We all know websites of Deep & Dark Web are not indexed, then how Dark Web Search Engines work?

According to Google’s online dictionary, the deep web is “the part of the World Wide Web that is not discoverable by means of standard search engines, including password-protected or dynamic pages and encrypted networks”. It is estimated that search engines like Google index only 4% of the entire world wide web, meaning that the deep web is nearly 25 times larger than the internet you and I have used our whole lives. Note: the deep web shouldn’t be confused with the “dark web”, which pertains strictly to pages containing illegal content such as child pornography, terrorist forums, and illegal auctions/transactions.

Google’s search engine functions by using “crawlers”. These crawlers start from a list of known web addresses, visit those pages, then follow the links contained on those pages, and continue following links found on the new pages, collecting information about each page as they go. Now, consider a single page in the deep web. Google’s search engine could be unable to find this page because of several reasons. For one, Google’s crawlers might never come across this page simply because no other previously crawled page links to it. Additionally, this page might require some sort of authentication such as filling out a search form and clicking submit, or having a certain certificate. Also, if a page contains illegal content, Google will likely not want that content appearing in search results, so they won’t index it. Finally, if the creator of a page doesn’t want it to be indexed by popular search engines, they can include a suitable robots.txt file, which tells the crawlers not to index the page. If the crawlers index the page anyways, then legal action can be taken against the creator of the crawlers, and the search engine can end up on a bot reporting site like <http://www.botreports.com/badbots/>.

Now for example, we consider Ahmia – a deep web search engine to understand how it searches the deep web or dark web. Ahmia essentially collects .onion URLs from the Tor network, then feeds these pages to their index provided that they don’t contain a robots.txt file saying not to index them. Additionally, Ahmia allows onion service operators to register their own URLs, enabling them to be found. Through continuously collecting .onion URLs, Ahmia has created one of the largest indexes of the deep web. But it is true that it still comes nowhere near to scratching the surface of the whole deep web, but it indexes a good portion of the content that most people would want to look for.

We can use another example such as Uncensored Hidden Wiki. It operates differently. Anyone can register on the Uncensored Hidden Wiki, and after that, anyone can edit the links contained in the database. The search engine operates by searching the provided descriptions of the pages at these links. This certainly has its pros and cons. On the bright side, crowd-sourcing the links is one of the best ways to collect a large number of useful URLs, and keep them up to date (especially since .onion domain names change extremely often). On the other hand, anyone can change the links to wherever they want, or alter the descriptions of the links. The negatives of this can be mitigated by site admins to ensure that the links are usually accurate, but there are no guarantees when using the links on this page. Additionally, the Uncensored Hidden Wiki has its name for a reason, as the content of that page is certainly uncensored.

While the “deep web search engines” mentioned above are capable of indexing a good part of the deep web, the vast majority of it remains unindexed, and no search engine is capable of finding everything contained in it. The best deep web search engines function in various ways, whether it be crowd-sourcing URLs and page descriptions or continuously collecting them, but they certainly do not function in similar ways to traditional search engines such as Google. If you want to learn more about the deep web, you can find plenty of information about the deep web using Google (how ironic). If you want to search the deep web yourself, here’s my advice: don’t. Especially if you’ve never heard of terms like .onion, Tor gateways, proxies, botnets, Trojans, etc. If you’re anything like me then you have no business searching the deep web, as it can be dangerous if you’re not extremely careful protecting your identity, even when using the search engines mentioned above in conjunction with the Tor browser. The deep web also has very little that you or I would find interesting, and plenty of things that neither you nor I want to see.

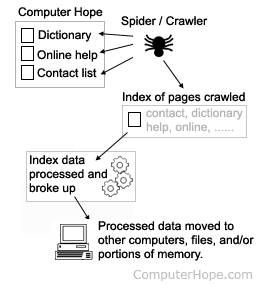
Surface Web Search Engine

Search engines are answer machines. They scour billions of pieces of content and evaluate thousands of factors to determine which content is most likely to answer your query.

A search engine is software, usually accessed on the Internet that searches a database of information according to the user's query. The engine provides a list of results that best match what the user is trying to find. Today, there are many different search engines available on the Internet, each with their own abilities and features. The first search engine ever developed is considered Archie, which was used to search for FTP files and the first text-based search engine is considered Veronica. Currently, the most popular and well-known search engine is Google. Other popular search engines include AOL, Ask.com, Baidu, Bing, and Yahoo.

**How a search engine works**

Because large search engines contain millions and sometimes billions of pages, many search engines not only search the pages but also display the results depending on their importance. This importance is commonly determined by using various algorithms.



As illustrated in the image on the right, the source of all search engine data is a spider or crawler, which automatically visits pages and indexes their contents.

Once a page is crawled, the data contained in the page is processed and indexed. Often, this can involve the steps below.

* Strip out stop words.
* Record the remaining words in the page and the frequency they occur.
* Record links to other pages.
* Record information about any images, audio, and embedded media on the page.

The data collected is used to rank each page. These rankings then determine which pages to show in the search results and in what order. Finally, once the data is processed, it is broken up into one or more files, moved to different computers, or loaded into memory where it can be accessed when a search is performed.

Search engines use proprietary algorithms to index and correlate data, so every search engine has its own approach to finding what you're trying to find. Its results may be based on where you're located, what else you've searched for, and what results were preferred by other users searching for the same thing, for example. Each search engines will weigh these factors in a unique way, and offer you different results.

Search engines have three primary functions:

1. **Crawl:** Scour the Internet for content, looking over the code/content for each URL they find.
2. **Index:** Store and organize the content found during the crawling process. Once a page is in the index, it’s in the running to be displayed as a result to relevant queries.
3. **Rank:** Provide the pieces of content that will best answer a searcher's query, which means that results are ordered by most relevant to least relevant.

**Crawling -** Crawling is the process by which search engines discover updated content on the web, such as new sites or pages, changes to existing sites, and dead links.

To do this, a search engine uses a program that can be referred to as a ‘crawler’, ‘bot’ or ‘spider’ (each search engine has its own type) which follows an algorithmic process to determine which sites to crawl and how often.

As a search engine’s crawler moves through your site it will also detect and record any links it finds on these pages and add them to a list that will be crawled later. This is how new content is discovered.

**Indexing -** Once a search engine processes each of the pages it crawls, it compiles a massive index of all the words it sees and their location on each page. It is essentially a database of billions of web pages.

This extracted content is then stored, with the information then organized and interpreted by the search engine’s algorithm to measure its importance compared to similar pages.

Servers based all around the world allow users to access these pages almost instantaneously. Storing and sorting this information requires significant space and both Microsoft and Google have over a million servers each.

Indexing is the process by which search engines organize information before a search to enable super-fast responses to queries. Searching through individual pages for keywords and topics would be a very slow process for search engines to identify relevant information. Instead, search engines (including Google) use an inverted index, also known as a reverse index.

**Reverse Index -** An inverted index is a system wherein a database of text elements is compiled along with pointers to the documents which contain those elements. Then, search engines use a process called tokenization to reduce words to their core meaning, thus reducing the amount of resources needed to store and retrieve data. This is a much faster approach than listing all known documents against all relevant keywords and characters.

**Cached version of A Page -** In addition to indexing pages, search engines may also store a highly compressed text-only version of a document including all HTML and metadata.

The cached document is the latest snapshot of the page that the search engine has seen.

The cached version of a page can be accessed (in Google) by clicking the little green arrow next to each search result’s URL and selecting the cached option. Alternatively, you can use the ‘cache:’ Google search operator to view the cached version of the page.

Bing offers the same facility to view the cached version of a page via a green down arrow next to each search result but doesn’t currently support the ‘cache:’ search operator.

**PageRank -** “PageRank” is a Google algorithm named after the co-founder of Google, Larry Page (yes, really!) It is a value for each page calculated by counting the number of links pointing at a page in order to determine the page’s value relative to every other page on the internet. The value passed by each individual link is based on the number and value of links which point to the page with the link.

PageRank is just one of the many signals used within the large Google ranking algorithm. An approximation of the PageRank values were initially provided by Google but they are no longer publicly visible.

While PageRank is a Google term, all commercial search engines calculate and use an equivalent link equity metric. Some SEO tools try to give an estimation of PageRank using their own logic and calculations. For example, Page Authority in Moz tools, TrustFlow in Majestic, or URL Rating in Ahrefs. DeepCrawl has a metric called DeepRank to measure the value of pages based on the internal links within a website.

**How PageRank flows through pages**

Pages pass PageRank, or link equity, through to other pages via links. When a page links to content elsewhere it is seen as a vote of confidence and trust, in that the content being linked to is being recommended as relevant and useful for users. The count of these links and the measure of how authoritative the linking website is, determines the relative PageRank of the linked-to page.

PageRank is equally divided across all discovered links on the page. For example, if your page has five links, each link would pass 20% of the page’s PageRank through each link to the target pages. Links which use the rel=”nofollow” attribute do not pass PageRank.

**How Search Engine Indexing Works (In Detail)**

Once the spiders have completed the task of finding information on Web pages (and we should note that this is a task that is never actually completed -- the constantly changing nature of the Web means that the spiders are always crawling), the search engine must store the information in a way that makes it useful. There are two key components involved in making the gathered data accessible to users:

* The information stored with the data
* The method by which the information is indexed

In the simplest case, a search engine could just store the word and the URL where it was found. In reality, this would make for an engine of limited use, since there would be no way of telling whether the word was used in an important or a trivial way on the page, whether the word was used once or many times or whether the page contained links to other pages containing the word. In other words, there would be no way of building the ranking list that tries to present the most useful pages at the top of the list of search results.

To make for more useful results, most search engines store more than just the word and URL. An engine might store the number of times that the word appears on a page. The engine might assign a weight to each entry, with increasing values assigned to words as they appear near the top of the document, in sub-headings, in links, in the Meta tags or in the title of the page. Each commercial search engine has a different formula for assigning weight to the words in its index. This is one of the reasons that a search for the same word on different search engines will produce different lists, with the pages presented in different orders.

Regardless of the precise combination of additional pieces of information stored by a search engine, the data will be encoded to save storage space. For example, the original Google paper describes using 2 bytes, of 8 bits each, to store information on weighting -- whether the word was capitalized, its font size, position, and other information to help in ranking the hit. Each factor might take up 2 or 3 bits within the 2-byte grouping (8 bits = 1 byte). As a result, a great deal of information can be stored in a very compact form. After the information is compacted, it's ready for indexing.

An index has a single purpose: It allows information to be found as quickly as possible. There are quite a few ways for an index to be built, but one of the most effective ways is to build a hash table. In hashing, a formula is applied to attach a numerical value to each word. The formula is designed to evenly distribute the entries across a predetermined number of divisions. This numerical distribution is different from the distribution of words across the alphabet, and that is the key to a hash table's effectiveness.

In English, there are some letters that begin many words, while others begin fewer. You'll find, for example, that the "M" section of the dictionary is much thicker than the "X" section. This inequity means that finding a word beginning with a very "popular" letter could take much longer than finding a word that begins with a less popular one. Hashing evens out the difference, and reduces the average time it takes to find an entry. It also separates the index from the actual entry. The hash table contains the hashed number along with a pointer to the actual data, which can be sorted in whichever way allows it to be stored most efficiently. The combination of efficient indexing and effective storage makes it possible to get results quickly, even when the user creates a complicated search.

**Indexing: How do search engines interpret and store your pages?**

Once you’ve ensured your site has been crawled, the next order of business is to make sure it can be indexed. That’s right — just because your site can be discovered and crawled by a search engine doesn’t necessarily mean that it will be stored in their index. In the previous section on crawling, we discussed how search engines discover your web pages. The index is where your discovered pages are stored. After a crawler finds a page, the search engine renders it just like a browser would. In the process of doing so, the search engine analyzes that page's contents. All of that information is stored in its index.

**Are pages ever removed from the index?**

Yes, pages can be removed from the index! Some of the main reasons why a URL might be removed include:

* The URL is returning a "not found" error (4XX) or server error (5XX) – This could be accidental (the page was moved and a 301 redirect was not set up) or intentional (the page was deleted and 404ed in order to get it removed from the index)
* The URL had a noindex Meta tag added – This tag can be added by site owners to instruct the search engine to omit the page from its index.
* The URL has been manually penalized for violating the search engine’s Webmaster Guidelines and, as a result, was removed from the index.
* The URL has been blocked from crawling with the addition of a password required before visitors can access the page.

If you believe that a page on your website that was previously in Google’s index is no longer showing up, you can use the URL Inspection tool to learn the status of the page, or use Fetch as Google which has a "Request Indexing" feature to submit individual URLs to the index. (Bonus: GSC’s “fetch” tool also has a “render” option that allows you to see if there are any issues with how Google is interpreting your page).

**Search Engine Ranking Process**

When someone performs a search, search engines scour their index for highly relevant content and then orders that content in the hopes of solving the searcher's query. This ordering of search results by relevance is known as ranking. In general, you can assume that the higher a website is ranked, the more relevant the search engine believes that site is to the query.

It’s possible to block search engine crawlers from part or all of your site, or instruct search engines to avoid storing certain pages in their index. While there can be reasons for doing this, if you want your content found by searchers, you have to first make sure it’s accessible to crawlers and is indexable. Otherwise, it’s as good as invisible.