

How Search Engine Works

Shaon Tewari

MCA Student, Department of Information Technology, Sikkim Manipal University, Haldia, India

Abstract: Internet is a way by which we can communicate from one Computer to other Computer.

Internet is the largest source of Information today. It offers information in large amounts on a variety of subjects. This very fact, in a way could hinder the use of the net. This is because; finding relevant information on the Net becomes very difficult. Over the years, with developments of the Net, a

number of tools to help find information have been designed initial search tools included Archie, Gopher, Veronica, WAIS and many other. With improvements in technology, search engines have also improved. Today a large number of powerful search engines with a variety of search strategies are available.

Everytime we go about in the internet, we most probably go to Google.com once in a while to look for something in the internet. That specific activity of looking for something is called search. When we type in something in Google.com's text area, what we are actually doing is asking Google's search engine to retrieve the information we just typed in – from their database back out to us.

A search engine is a long series and combination of codes which when activated, goes out and does three main things: Index, Retrieve and Rank. Those are the three major functions of a search engine – don't worry, I shall be discuss them to give a more in-depth understanding of how a search engine works.

Now a days, we don't necessarily have to submit our website to the search engines, they'll automatically come looking for us as they crawl the web. Crawling by the way is the term used when a search engine reads the codes of a website. Always keep in mind that All websites are made up of codes. What we see is only an end-product of those codes when read by an HTML browser.

When we enter the words we are looking for in Google.com's text area and we hit the enter button, what we just did is called a search query. And what Google will do next is look up in it's database for any related keyword to our query and retrieve it for us to display in its SERP or Search Engine Results Page.

So that's it for this lesson. We'll discuss more about how search engines work in the next lessons.

Keywords: search engine

1. History of search engine

The need for search engines was first noted in 1945 when American engineer and scientist Vannevar Bush published an article in The Atlantic Monthly, emphasizing the necessity for an expansive index for all knowledge: "[Information] has been extended far beyond our present ability to make real use of the record. A record, if it is to be useful to science, must be continuously extended, it must be stored...Our ineptitude in getting at the record is largely caused by the artificiality of the systems of indexing. The human mind does not work this way. It operates by association."

Decades later, college students and electrical engineers attempted to make this kind of index a reality. The development of Archie (named for "archives"), the very first search engine created in 1990, was designed to search and store directory listings on file transfer protocol sites. Similar engines began to follow suit, when Gopher (a site that indexed text files) came on the scene, it was followed by Veronica and Jughead, both of which corresponded to Gopher. These early versions paved the way for the technologies utilized by main line search engines today.

For example, JumpStation was the first world wide web resource-discovery tool which employed crawling and indexing. This was followed by the first full-text search engine, WebCrawler created in 1994 by Brian Pinkerton. Prior to this, only webpage names/domains were indexed in catalogues. As the internet expanded, search engines and web directory technologies increased, creating competitive platforms for the strongest emerging technologies to arise.

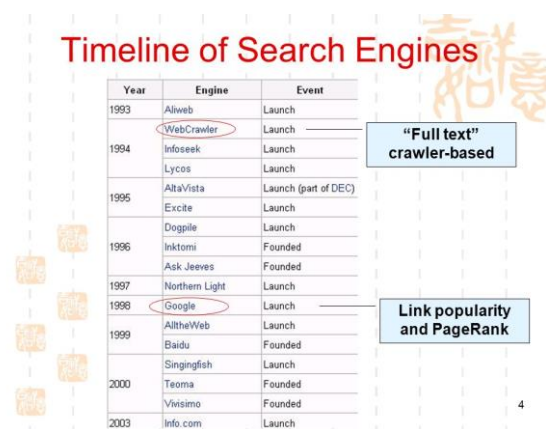


Fig. 1. Timeline of search engines

Search Engines Development Time Period:

- Archie (1990) was the first tool created by Alan Emtage and L. Peter Deutsch for indexing, and is considered the first basic search engine. What began as school project at McGill University in Montreal, was an index that predated the world wide web(WWW). Gopher, released in 1991 by students from the University of Minnesota, was a protocol used to index and search for documents online as a form of anonymous FTP. Archie, Gopher and similar counterparts lost traction in the late 90's.
- Lycos (1993) was created as a university project, but was

the first to attain commercial search engine success. In 1999 Lycos was the most visited search engine in the world and was available in 40 countries. Now it is currently comprised of a social network with email, webhosting and media entertainment pages.

- Yahoo! (1994) started at Stanford University by Jerry Yang and David Filo that became a web portal and search engine.
- WebCrawler (1994) created by Brian Pinkerton WebCrawler was the first crawler which indexed complete pages online. AOL purchased WebCrawler, using the technology for their network, and when Excite purchased WebCrawler, AOL used Excite to run their program NetFind. WebCrawler was one of the foundational search engines.
- AltaVista (1995) an industry leader, was once the most popular search engines of its time. It differed from its contemporaries because of two factors: Alta Vista used a multi-threaded crawler (Scooter) that covered more webpages than people knew existed at the time. It also had a well-organized search-running back-end advanced hardware. By 1996 AltaVista had become the sole search results provider for Yahoo. In 2003 Alta Vista was bought by Overture Services, Inc., which only months later was acquired by Yahoo.
- Looksmart (1995) competed with Yahoo! Directory and had an initial a goal to create a substantial directory of websites. When it went public in 1999, it lost a fair amount of customers. By 2002 Looksmart became a pay-per-click provider, and after being dropped by Microsoft, bought a search engine called WiseNut. Sadly it never gained serious traction, and Looksmart lost its momentum.
- WiseNut (2001) was a crawler-based search engine that was introduced as a beta, and was owned by Looksmart. Initially the site was well reputed with an unsullied database, and an automatic clustering of search results by using a technology called WiseGuide. Looksmart bought WiseNut in 2002, and was eventually closed in 2007.
- Excite (1995) Founded originally as "Architext" by Stanford University students, Excite was launched officially having purchased two search engines (Magellan and WebCrawler), and signed exclusive agreements with Microsoft and Apple.
- Hotbot (1996) a search engine also popular in the 90's was launched by Wired Magazine, and is now owned by Lycos.
- Dogpile (1996) was a search engine developed by Aaron Flin and shortly thereafter sold to Go2net. Now Dogpile fetches results from Google, Yahoo, and Yandex.
- Google (1996) Started for a research project by Stanford students Larry Page and Sergey Brin. They created a search engine that would rank websites based on the number of other websites that linked to that page. Prior to this, other engines have ranked sites based on the number of times the search term appeared on the webpage. This strategy developed the world's most successful search engine today.
- MSN Search (1998) was the engine used by Microsoft, sourcing search results from Inktomi, and later Looksmart. By 2006 Microsoft started performing their own image searches, and MSN became branded as Windows Live Search, then Live Search, and finally to Bing (2009) which was set to replace the Yahoo search engine.
- ASK (1996) was originally titled "AskJeeves.com" and was designed by Garret Gruener and David Warthen in Berkeley, CA. The goal was to provide users with answers to queries typed with normal every-day language and colloquialisms. It was acquired in 2005 by IAC and continues to grow with over 100 million users.
- Teoma (2000) meaning "expert" in Gaelic, was a search engine created by professor Apostolos Gerasoulis and Tao Yang at Rutgers University. Teoma's subject-specific technology centered on a link popularity algorithm which allowed pages to rank higher if other pages with a similar content and subject matter linked back to the page. Teoma was acquired by Ask Jeeves in 2001, and rebranded as Ask.com.
- Infoseek (1994) was a search engine begun by Steve Kirsch, and was bought by The Walt Disney Company in 1998, merging with Starwave to become go.com. Eventually it was replaced by Yahoo, and no longer exists.
- Overture (1998) was originally named "GoTo," where top listings were sold on a cost-per-click or pay-per-click basis. In 2000, they began driving traffic by placing its paid listings on other search engines, and was eventually bought by Yahoo in 2003.
- Alltheweb (1999) began in 1994 out of FTP Search, from Norwegian University of Science and Technology, when then turned into Fast Search & Transfer, or FAST. Alltheweb (1999) was said to have once rivaled Google, but the number of users declined when Overture bought the company in 2003.
- AOL Search (1999) bought Web Crawler (one of the major crawler-based engines of it's time) in 1995, and after a number of deals, purchases and exchanges, AOL relaunched their search engine, calling it AOL Search. Teaming with Google, the search engine relaunched in 2006 with newer features including video, search marketplace, etc.
- Year-2000:- Baidu, a Chinese company that would grow to provide many search-related services, launches.
- Year 2002-2003:- Yahoo! buys Inktomi (2002) and then Overture Services Inc. (2003) which has already bought AlltheWeb and Altavista. Starting 2003, Yahoo! starts using its own Yahoo Slurp web crawler to power Yahoo! Search. Yahoo! Search combines the technologies of all Yahoo!'s acquisitions (until 2002, Yahoo! had been using Google to power its search).
- Year-2004:- Microsoft starts using its own indexer and crawler for MSN Search rather than using blended results from LookSmart and Inktomi.
- Year-2005: Overture Services Inc. owner Bill Gross launches the Snap search engine, with many features such

as display of search volumes and other information, as well as sophisticated auto-completion and related terms display. It is unable to get traction and soon goes out of business.

- Year-2006:-2009:- Wikia launches Wikia Search, a search engine based on human curation, but then shuts it down. Relevant dates: publicly proposed December 23, 2006 and January 31, 2007, private pre-alpha December 24, 2007, toolbar release August 2008, shutdown March–May 2009.
- Year-2010:- Google launches Google Instant, described as a search-before-you-type feature: as users are typing, Google predicts the user's whole search query (using the same technology as in Google Suggest, later called the auto complete feature) and instantaneously shows results for the top prediction.
- Year-2011: Google, Yahoo!, and Microsoft announce Schema.org, a joint initiative that supports a richer range of tags that websites can use to convey better information.
- Year-2012: Google launches Search plus Your World, a deep integration of one's social data into search.

Definition & Components of Search Engine

Search Engine refers to a huge database of internet resources such as web pages, newsgroups, programs, images etc. It helps to locate information on World Wide Web. Or in other words, A search engine is an information retrieval system designed to help find information stored on a computer system. The search results are usually presented in a list and are commonly called hits.

Search engine is a service that allows Internet users to search for content via the World Wide Web (WWW). A user enters keywords or key phrases into a search engine and receives a list of Web content results in the form of websites, images, videos or other online data. The list of content returned via a search engine to a user is known as a Search Engine Results Page (SERP).

User can search for any information by passing query in form of keywords or phrase. It then searches for relevant information in its database and return to the user.

Search Engine Components

Generally, there are three basic components of a search engine as listed below:

1. Web Crawler
2. Database
3. Search Interfaces

1. **Web crawler:** It is also known as spider or bots. It is a software component that traverses the web to gather information.
2. **Database:** All the information on the web is stored in database. It consists of huge web resources.
3. **Search Interfaces:** This component is an interface between user and the database. It helps the user to search through the database.

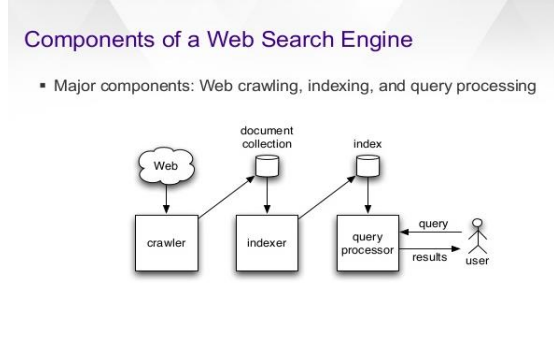


Fig. 2. Components of a Search Engine

- **What are the main uses of web search engines?**
 Generally, people use search engines for one of three things: research, shopping, or entertainment. These are discussed below:

- **Using search engines for research:**

Most people who are using a search engine are doing it for research purposes. They are generally looking for answers or at least to data with which to make a decision. They're looking to find a site to fulfill a specific purpose.

- **Using search engines to shop:**

A smaller percentage of people, but still very many, use a search engine in order to shop. After the research cycle is over, search queries change to terms that reflect a buying mindset. Terms like “best price” and “free shipping” signal a searcher in need of a point of purchase.

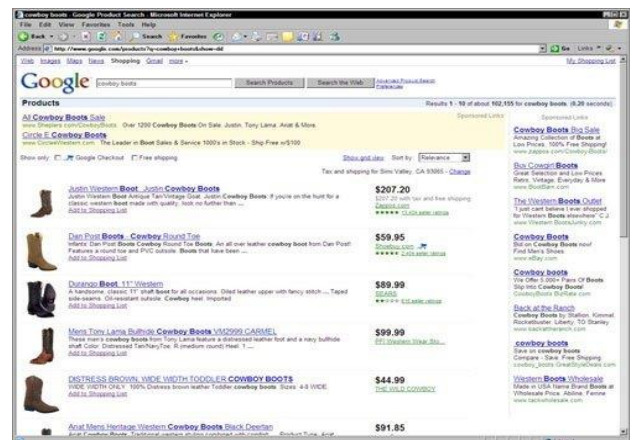


Fig. 3. Google search

- **Using search engines to find entertainment**

Research and shopping aren't the only reasons to visit a search engine. There are users out there who use the search engines as a means of entertaining themselves. They look up things like videos, movie trailers, games, and social networking sites. Technically, it's also research, but it's research used strictly for entertainment purposes.

2. Some Popular Search Engines and Their Working

Which are the 10 best and most popular search engines in the World? Besides Google and Bing there are other search engines that may not be so well known but still serve millions of search queries per day.

It may be a shocking surprise for many people but Google is not the only search engine available on the Internet today! In fact, there are a number of search engines that want to take Google's throne but none of them is ready (yet) to even pose a threat.



Fig. 4. List of Top 5 Most Popular Search Engines in the World (Updated 2019)

- *Google:*

It is one of the most important Search Engine. I discuss this search engine at previous. So no need for further introductions. The search engine giant holds the first place in search with a stunning difference of 65% from second in place Bing.

According to the latest netmarketshare report (November 2018) 73% of searches were powered by Google and only 7.91% by Bing.

Google is also dominating the mobile or tablet search engine market share with 81%.

- *Bing:*

Bing is Microsoft's attempt to challenge Google in search, but despite their efforts they still did not manage to convince users that their search engine can produce better results than Google.

Their search engine market share is constantly below 10%, even though Bing is the default search engine on Windows PCs.

- *Yahoo:*

Yahoo is one of the most popular email providers and holds the fourth place in search with 3.90% market share.

From October 2011 to October 2015, Yahoo search was powered exclusively by Bing. Since October 2015 Yahoo agreed with Google to provide search-related services and since then the results of Yahoo are powered both by Google and Bing. Yahoo is also the default search engine for Firefox browsers in the United States (since 2014).

- *Ask.com:*

It also has the general search functionality but the results returned lack quality compared to Google or even Bing and Yahoo.

What is Meta Search Engine & Their Types

A Meta search engine, otherwise it is known as an aggregator, is a search engine that sends queries to several search engines and either aggregates the results into one master list or categorizes the results by the search engines they come from. There are dozens of meta search engines across the Internet, and Dogpile is one main example.

At last, a meta search engine allows a user to enter a single query and field results from several sources. The idea is that this breadth of information allows users to get the best answers as quickly as possible.

Which Meta Search engines should we use for what?

- *Dogpile* — When we want a compilation of results from all the major search engines.
- *Ithaki* — When we want our aggregated results to be ranked by an internal algorithm.
- *Polymeta* — When we want a super intelligent meta search engine that can recognize colloquial language and uses an algorithm to organize results.
- *SearchSalad* — When we want a combination of search engine results and information from all the major review websites.
- *Seekz* — When we want a meta search engine that produces the most relevant results and removes all duplicates.

Advantages and Disadvantages of Search Engine

Although, Search Engine is most essential for the user to search any information from the website, But there are some common Advantages and Disadvantage are also given. Some of them are given below:

Advantages:

Search Engine has various advantages, Some of them are given below:

- The indexes of search engines are usually vast, representing significant portions of the Internet, offering a wide variety and quantity of information resources.
- The growing sophistication of search engine software enables us to precisely describe the information that we seek.
- The large number and variety of search engines enriches the Internet, making it at least appear to be organized.
- Search engines scan the entire Web and keep comprehensive data on every page they catalog.
- In addition to keywords, search engines let you use advanced search options to refine your results. These options help make your searches more flexible and sophisticated.
- Some search engines, such as LexisNexis, specialize in legal or other specialized, scholarly information; these sites charge a fee to use their services.

Disadvantages:

Search Engine has various dis-advantages, Some of them are given below:

- They do not crawl the web in “real time”.
- If a site is not linked or submitted it may not be accessible.
- Not every page of a site is searchable.
- Special tools needed for the Invisible/Deep Web.
- Few search engines search the full text of Web pages.
- There are so many external factors around competition and search trends, and we will never know 100% how each search engine's algorithms work. Therefore, it's very difficult to ensure results.
- They filter results according to the information that a particular user has given them, which rarely provides an accurate reflection of a user’s real interests.
- Rankings are almost certainly affected in some way by the search engine companies’ commercial interests: their dependance on paid advertising, and their promotion of their own products.

Search Engines Architecture:

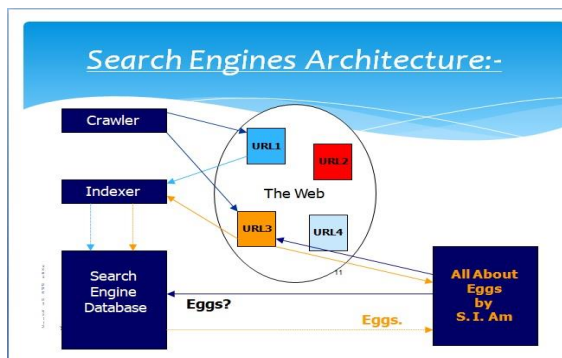


Fig. 5. Search engines architecture

How Search Engine Works?

- Search engines actually search the web to “index” all of the words on the Internet. This is a huge job. A good search engine will be able to find the exact words that we are looking for.
- Some search engines search only certain types of sites, like medical sites, animal sites, and so on. The results they give will be more limited, but probably more through.
- All search engines work using a 3 phase approach to managing , ranking and returning search results. But a lot of people have no idea what is happening behind that search box when they type in their search queries. So just how do Google, Bing and the rest of them work out what is on the web, what is relevant to our general query and which specific

websites should be ranked highly?

- There are three functions which need to be done:

Steps Followed by a Search Engine to Display Results:

Every Search engines uses their own software programs to display results, but the way they work is quite similar. Every search engine performs these tasks to increase their database and show the proper and relevent results to any search query.

- Crawling
- Indexing
- Ranking
- Results

• *Spider or Crawler:*

1. The spider visits a web page, reads it, and then follows links to other pages within the site.
2. This is what it means when someone refers to a site being "spidered" or "crawled".
3. This is also known as “harvesting”.
4. The spider returns to the site on a regular basis, such as every month or two, to look for changes._Crawling the Web, following links to find pages.

Indexer:

1. Everything the spider finds goes into the second part of a search engine, the *index*.
2. The index, sometimes called the catalog, is like a giant book containing a copy of every web page that the spider finds.
3. If a web page changes, then this book is updated new information.
4. *Indexing* the pages to create an index from every word to every place it occurs.

Ranking:

Now the search engine have a list of websites or WebPages for different queries, but here one more bigger issue comes in the action. Now search engine has to decide which website is more relevant to the search query.

Most of the search engines uses their secret algorithms to rank the webpages, google alsos have the same. They check more than 200 signals to decide the rank of a webpage which includes how is the title, meta description has been written. How many other webpages are linking to that page and how fast a website loads.

Google always try to show the best result to their user or for every search query, and this is the sole purpose and mission of google to show the best possible result to their users.

Result:

Most of the Search engines examine all the pages on the World Wide Web which are allowed to crawl, categories them and put them into a logical order according to their parameters and put a list of results in front of you when you search for something.

How to Search?

Most search engines build an index based on crawling, which is the process through which engines like Google, Yahoo and others find new pages to index. Mechanisms known as bots or

spiders crawl the Web looking for new pages (1). The bots typically start with a list of website URLs determined from previous crawls. When they detects new links on these pages, through tags like HREF and SRC, they add these to the list of sites to index. Then, search engines use their algorithms to provide you with a ranked list from their index of what pages you should be most interested in based on the search terms you used.

Then, the engine will return a list of Web results ranked using its specific algorithm. On Google, other elements like personalized and universal results may also change your page ranking. In personalized results, the search engine utilizes additional information it knows about the user to return results that are directly catered to their interests. Universal search results combine video, images and Google news to create a bigger picture result, which can mean greater competition from other websites for the same keywords.

- Search by word: rainforest
- Search by phrase: rainforest destruction
- Search by question: Why are rainforests being destroyed?
- Search by name of an author: Steve Knight
- Google.com

When searching, keep these things in mind:

- Capital letters or lower case letters do not make any difference in search results.
- If you put your words in “ “ the search engines will search for that exact phrase only
- If you do not put your words in “ “ the search engines will search for pages that include all of those words. They may not be all together.

Here are the top elements to edit when designing our store for Search Engine:

- *Architecture* - Make websites that search engines can crawl easily. This includes several elements, like how the content is organized and categorized and how individual websites link to one another. An XML sitemap can allow you to give a list of URLs to search engines for crawling and indexing.
- *Content* - Great content is one the most important elements for SEO because it tells search engines that your website is relevant. This goes beyond just keywords to writing engaging content your customers will be interested in on a frequent basis.
- *Links* - When a lot of people link to a certain site, that alerts search engines that this particular website is an authority, which makes its rank increase. This includes links from social media sources. When your site links to other reputable platforms, search engines are more likely to rate your content as quality also.
- *Keywords*- The keywords you use are one of the primary methods search engines use to rank you. Using carefully selected keywords can help the right

customers find you. If you run a jewelry store but never mention the word "jewelry," "necklace," or "bracelet," Google's algorithm may not consider you an expert on the topic.

- *Title descriptions* - While it may not show up on the website, search engines do pay attention to the title tag in your site's html code, the words between < title > < /title >, because it likely describes what the website is about, like the title of a book or a newspaper headline.
- *Page content* - Don't bury important information inside Flash and media elements like video. Search engines can't see images and video or crawl through content in Flash and Java plugins.
- *Internal links* - Including internal links helps search engines crawl your website more effectively, but also boosts what many SEO professionals refer to as "link juice." In other words, it has the same benefit of any link to our site: It demonstrates the value of our content.

Why Search Fail:

Search was fail due to some important reason. Some of them are given below:

- *Empty search:* If we keep the search box is empty in the Internet browser, then the search was fail.
- *Nothing on the site on that topic (scope):* If we type odd topic in this search box then the search may be fail to search that we want to given.
- *Misspelling or typing mistakes:* When we type wrong spelling of the word in the search box then search may be fail.
- *Vocabulary differences:* It is the most important reason for failing search. The vocabulary difference between the different words can fail to search that we want to given.
- *Restrictive search defaults:* Some topics are present which was restricted by the search engine. If we type these topics in our search box then search may be failed.
- *Restrictive search choices:* It is the most another important reason for search failing. If our choice is restricted by the search engine then the search was fail.
- *Software failure:* If our browsing software like-Google chrome, Firefox are outdated, then we need to update our browser regularly to avoid search fail.

Do all Search Engines gives the same results?

- No, they do not. Not even close.
- A Meta search engine should give the most results.
- Google gives very comprehensive results, quickly.
- The type of device used for the search (desktop, laptop, phone, tablet).
- Our personal search history.

- Whether we are logged in to a Google Account while searching.
- Our Geographic location.
- What type of browser we are using.
- The number of Google-generated ads on the page.
- What type of search we are doing.

Search Results Interface:

What users see after they click the Search button?

- The most visible part of search
- Elements of the results page
- Page layout and navigation
- Results header
- List of results items
- Results footer

How Google Works:

Google's search engine is a powerful tool. Without search engines like Google, it would be practically impossible to find the information you need when you browse the Web. Like all search engines, Google uses a special [algorithm](#) to generate search results. While Google shares general facts about its algorithm, the specifics are a company secret. This helps Google remain competitive with other search engines on the Web and reduces the chance of someone finding out how to abuse the system.

Google uses automated programs called spiders or crawlers, just like most search engines. Also like other search engines, Google has a large index of keywords and where those words can be found. What sets Google apart is how it ranks search results, which in turn determines the order Google displays results on its search engine results page (SERP). Google uses a trademarked algorithm called Page Rank, which assigns each Web page a relevancy score.

A Web page's PageRank depends on a few factors:

- *The frequency and location of keywords within the Web page:* If the keyword only appears once within the body of a page, it will receive a low score for that keyword.
- *How long the Web page has existed:* People create new Web pages every day, and not all of them stick around for long. Google places more value on pages with an established history.
- *The number of other Web pages that link to the page in question:* Google looks at how many Web pages link to a particular site to determine its relevance.
- How do search engines crawl the web?

Google's first job is to 'crawl' the web with 'spiders.'

These are little-automated programs or bots that scour the 'net for any and all new information.

The spiders will take notes on your website, from the titles you use to the text on each page to learn more about who you are, what you do, and who might be interested in finding you.

How Does Google Search Work in 2018?

Since the very beginning, Google has focused on directing its

users to high quality, relevant results. And its algorithm updates reflect that focus.

So while it can be difficult to predict exactly which changes are on the horizon for the SEO industry, it's safe to assume that they'll have the same basic goal of providing users with the best possible results.

Of course, that doesn't mean you don't need to pay attention when the search engine announces a new update.

While some of these updates may be aimed at penalizing SEOs using unscrupulous tactics, others will give you insight into the practices you should be using on your site.

For example, Google's more recent updates have focused on adapting to search trends, prioritizing user security, and making it easier for users to access information directly on results pages.

There are hundreds of millions of WebPages available over the Web, most of which are titled on the whim of their author. Let's see how Google is able to offer incredible search results out of hundreds of millions of webpages available.

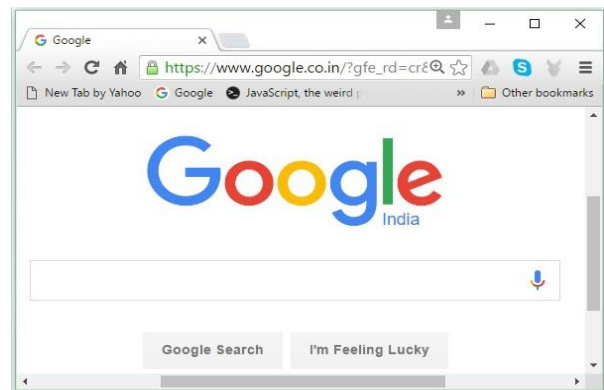


Fig. 6. Google Search Engine

How Google works:

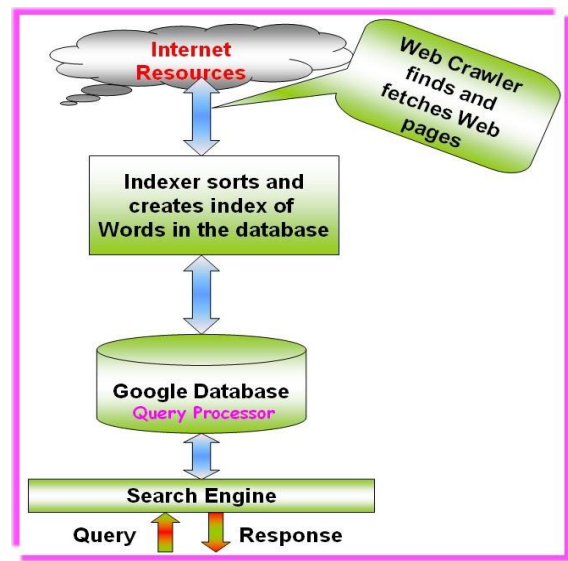


Fig. 7. How google works

- Here I shall know how Google creates the index and the database of documents that it accesses when processing a query.
- Google runs on a distributed network of thousands of low-cost computers and can therefore carry out fast parallel processing.
- Parallel processing is a method of computation in which many calculations can be performed simultaneously, significantly speeding up data processing.
- The general appearance of Google home page is shown in Fig. 6. Google has a regular tendency to amaze its users. So quite often it stages unique logos at its home page to honour certain events.

Google has three distinct parts:

- *Google bot*, a web crawler that finds and fetches web pages.
- *The indexer*, that sorts every word on every page and stores the resulting index of words in a huge database.
- *The query processor*, which compares your search query to the index and recommends the documents that it considers most relevant.

1. Google bots

- Google bots is Google's web crawling robot.
- which finds and retrieves pages on the web and hands them off to the Google index
- Google bots consists of many computers requesting and fetching pages much more quickly than we can with our web browser.
- In fact, Google bot can request thousands of different pages simultaneously.

2. Google's Indexer

- Googlebot gives the indexer the full text of the pages it finds.
- These pages are stored in Google's index database.
- This index is sorted alphabetically by search term, with each index entry storing a list of documents in which the term appears and the location within the text where it occurs.
- This data structure allows rapid access to documents that contain user query terms.

3. Google's Query Processor

- The query processor has several parts, including the user interface (search box), the "engine" that evaluates queries and matches them to relevant documents, and the results formatter.
- PageRank is Google's system for ranking web pages.
- A page with a higher PageRank is deemed more important and is more likely to be listed above a page with a lower PageRank.

How do we make our searches better?

These days, everyone is expected to be up to speed on Internet search techniques. But there are still a few tricks that some users may apply: This is given below:

Google has been fanatical about speed. There is little doubt that it has built an incredibly fast and thorough search engine. Unfortunately, the human element of the Internet search equation is often overlooked. These 7 tips are designed to improve that human element and better our Internet search skills.

1. Use unique, specific terms

It is simply amazing how many Web pages are returned when performing a search. We might guess that the terms blue dolphin are relatively specialized. A Google search of those terms returned 2,440,000 results! To reduce the number of pages returned, use *unique* terms that are *specific* to the subject you are researching.

2. Use the minus operator (-) to narrow the search

How many times have you searched for a term and had the search engine return something totally unexpected? Terms with multiple meanings can return a lot of unwanted results. The rarely used but powerful minus operator, equivalent to a Boolean NOT, can remove many unwanted results. For example, when searching for the insect caterpillar, references to the company Caterpillar, Inc. will also be returned. Use Caterpillar -Inc to exclude references to the company or Caterpillar -Inc -Cat to further refine the search.

3. Use quotation marks for exact phrases

I often remember parts of phrases I have seen on a Web page or part of a quotation I want to track down. Using quotation marks around a phrase will return only those exact words in that order. It's one of the best ways to limit the pages returned. Example: "Be nice to nerds". Of course, you must have the phrase exactly right — and if your memory is as good as mine, that can be problematic.

4. Don't use common words and punctuation

Common terms like *a* and *the* are called stop words and are usually ignored. Punctuation is also typically ignored. But there are exceptions. Common words and punctuation marks *should* be used when searching for a specific phrase inside quotes. There are cases when common words like *the* are significant. For instance, Raven and The Raven return entirely different results.

5. Capitalization

Most search engines do not distinguish between uppercase and lowercase, even within quotation marks. The following are all equivalent:

- technology
- Technology
- TECHNOLOGY
- "technology"
- "Technology"

6. Drop the suffixes

It's usually best to enter the base word so that you don't

exclude relevant pages. For example, bird and not birds, walk and not walked. One exception is if you are looking for sites that focus on the act of walking, enter the whole term walking.

7. Maximize Auto Complete

Ordering search terms from general to specific in the search box will display helpful results in a drop-down list and is the most efficient way to use AutoComplete. Selecting the appropriate item as it appears will save time typing. We have several choices for how the AutoComplete feature works.

8. Use a colon to search specific sites

There may be an instance where you need to Google search for articles or content on a certain website. The syntax is very simple and we'll show you below.

- *Sidney Crosby site:nhl.com*

This will search for all content about famous hockey player Sidney Crosby, but only on NHL.com. All other search results will be removed. If you need to find specific content on a particular site, this is the shortcut you can use.

3. Conclusion

Search engines offer users vast and impressive amounts of information, available with a speed and convenience few people could have imagined one decade ago. Their capabilities are expanding practically by the day. Soon it will seem routine to be able to search the contents of vast libraries of books; to find selected portions of video streams or audio recordings; to benefit from personalized searches that remember a user's preferences and keep track of changing geographical locations. Audio searching and search results will be available for the blind; "implicit searching" will anticipate users' queries and have answers ready.

Today's internet users are very positive about what search engines already do, and they feel good about their experiences when searching the internet. They say they are comfortable and confident as searchers and are satisfied with the results they find. They trust search engines to be fair and unbiased in returning results. And yet, people know little about how engines operate, or about the financial tensions that play into how engines perform their searches and how they present their search results. Furthermore, searchers largely don't notice or understand or discern the different kinds of search results that are being served up to them.

This odd situation, in which a growing population of users relies on technology most of them don't understand, highlights

the responsibility placed on search engine companies. They are businesses, in many cases extremely successful ones – but their effects on society are far more than merely commercial. One unexpected implication of our study is that search engines are attaining the status of other institutions – legal, medical, educational, governmental, and journalistic – whose performance the public judges by unusually high standards, because the public is unusually reliant on them for principled performance.

Now that we all have finished our search engine experimentation we have hopefully learned and accomplished much. Combining data and communicating with our colleagues should have been interesting. We have undoubtedly found worthwhile search engines, strategies and methods. This should make it easier to connect our own investigations and experiments to larger science ideas. Maybe we have even discovered some helpful websites in the process of our experiment. Congratulations and thank you all.

References

- [1] https://www.google.com/search?q=timeline+of+search+engines&source=lnms&tbm=isch&sa=X&ved=0ahUKEwi2nNXv6r3gAhVFFH0KHR0eAnUQ_AUIDigB&biw=1366&bih=664#imgrc=fYWPodBviFerHM
- [2] https://en.wikipedia.org/wiki/Web_search_engine
- [3] <https://www.the-reference.com/en/expertise/digital-marketing/search-engine-optimisation>
- [4] www.google.com
- [5] https://www.tutorialspoint.com/internet_technologies/search_engines.htm
- [6] <https://www.lifewire.com/web-search-tricks-to-know-4046148>
- [7] The Cyberspace Handbook by Jason Whittaker Routledge, 2004.
- [8] So Much Information, So Little Time: Evaluating Web Resources with Search Engines by Killmer, Kimberly A.; Koppel, Nicole B T H E Journal (Technological Horizons in Education), Vol. 30, No. 1, August 2002
- [9] Search Engine Technology Impetus for the Knowledge Revolution in Business Education by Hall, Owen P., Jr Journal of Interactive Learning Research, Vol. 15, No. 2, Summer 2004.
- [10] <http://www.oxfordreference.com/view/10.1093/oi/authority.2011080310450343>
- [11] <https://www.techopedia.com/definition/12708/search-engine-world-wide-web>
- [12] <https://moz.com/beginners-guide-to-seo/how-search-engines-operate>
- [13] Web Search Savvy: Strategies and Shortcuts for Online Research by Barbara G. Friedman Lawrence Erlbaum Associates, 2005
- [14] So Much Information, So Little Time: Evaluating Web Resources with Search Engines by Killmer, Kimberly A.; Koppel, Nicole B, Journal (Technological Horizons in Education), Vol. 30, No. 1, August 2002.
- [15] 15. Search Engine Technology Impetus for the Knowledge Revolution in Business Education by Hall, Owen P., Jr Journal of Interactive Learning Research, Vol. 15, No. 2, Summer 2004.