**How Do Search Engines Work?**



 Mark As Completed Interactive Mode Discussion

In this lesson, we will learn about how search engines work, with a focus on following key points:

1. What are search engines?
2. What are the inner workings of search engines?
3. What is search engine optimization (SEO)?

Imagine an internet without Google. If this existed, you wouldn't be able to access the web with ease by typing in some keywords related to your search. You would have to know all the URLs of websites that contained the information that you want.

Accessing the internet in such a case would become very difficult. Google is an example of a search engine, which has become a crucial gateway to access the internet.

What are Search Engines?

Let's discuss the formal definition of a search engine. Search engines are web-based tools, which allow users to search content on the internet according to input keywords and find webpages that contain content related to the search.

**Google** is the most popular search engine. However, it is not the only one. Yahoo, Bing, DuckDuckGo, and Yandex are among some other search engines.

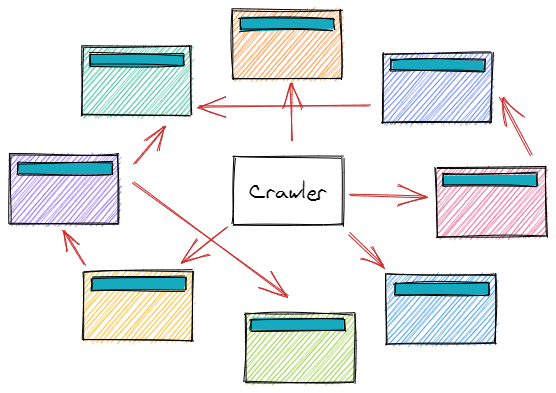
It is important to know how these search engines work. Suppose you create a website and want to share it with a larger audience, or want Google to display it in one of its top searches. How do you do that? You must be able to understand how search engines work, organize and analyze content for this purpose.

Internals of Search Engines

To work effectively, search engines must understand the kind of information that is requested, and what kind of webpages from the web are relevant to the requested information. For this purpose, search engines perform three important steps; **crawling**, **indexing**, and **ranking**.

Crawling

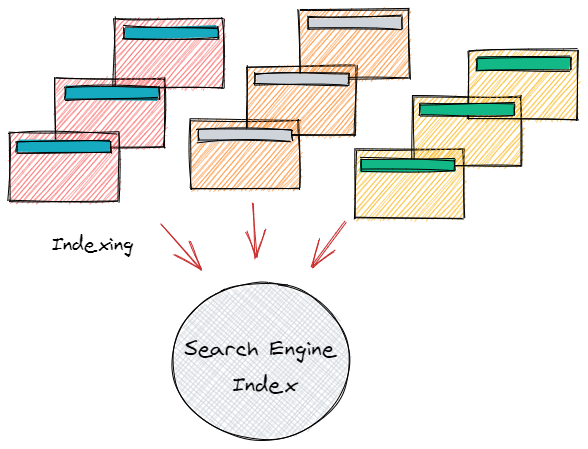
Search engines utilize **web crawlers**, which are bots that search the entire web for new, updated, and modified content. These web crawlers look for URLs, sitemaps (a file that provides information about content and pages on a website), or code of the website to find them. These may go further onto links that further link to another webpage. This leads to the discovery of new webpages. It is very similar to how you browse the internet for a web search, but here bots browse the information instead of humans!



Indexing

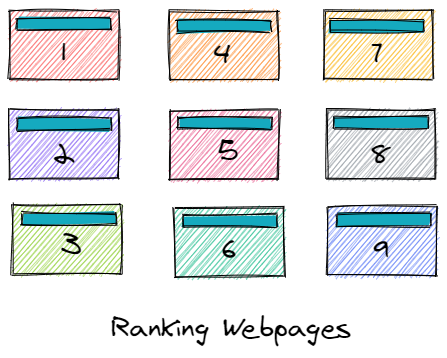
After crawling webpages, the bots that search the web will organize information through the process of indexing.

In indexing, search engines collect, analyze and store everything found in data structures called **indices**. Modifications, recent activity, keywords, and other significant features of the webpage are all stored in the index. Imagine an index as a library of books, but containing webpages instead of books!



Ranking

As the information regarding new webpages is organized in a search engine's index, they are then ranked according to content quality, currentness, and relevance. The ranking process involves the use of **search algorithms** which influence the decision regarding the information that would be displayed to the user from the search engine index. Different search engines use different search algorithms to perform the ranking of webpages.



**Search algorithms** employ techniques like artificial intelligence, machine learning, and deep learning to determine relevant results to search engine queries. These algorithms are updated from time to time to improve the search quality. They analyze features like term frequencies, website code, content relevance using data mining.

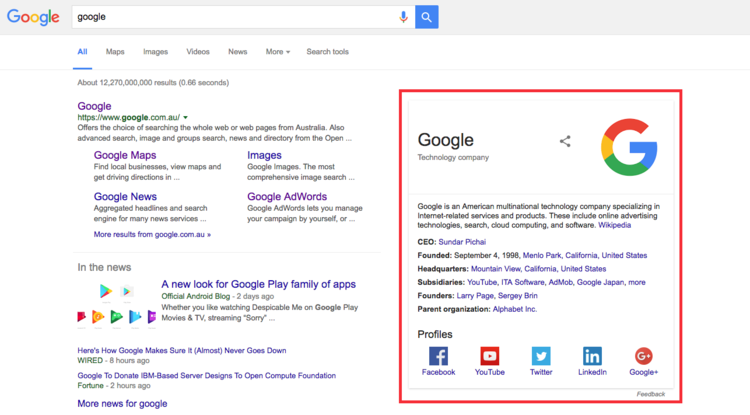
At a high level, search engines perform the above three actions whenever there is a search query. However, there may be other small steps involved during each of these actions.

Understanding Google's Search Engine

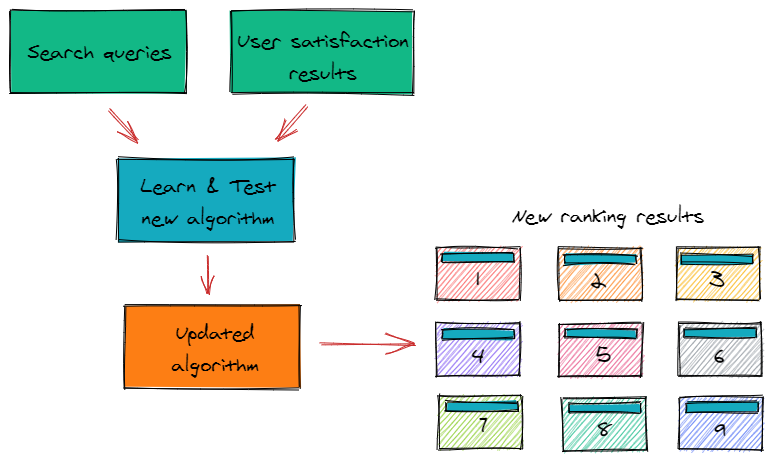
Let's understand the internals of search engines using an example of our favorite search engine, Google.

Google, like every other search engine, uses web crawlers that find new and updated websites using data from previous crawls or sitemaps. It analyzes text and visual content, along with the overall layout of the site, and decides if the particular site should appear in search results or not. Google also provides a tool, Search Console that allows site owners to view Google search traffic on their site, fix indexing problems, and note down tips on how they should further improve the content and layout site to make it more visible.

For indexing, Google uses a Search Index to store key features (keywords, freshness, among other things) of webpages that are searched after web crawling. Google has taken one step further in the indexing process and has introduced a knowledge graph (illustrated below), a powerful feature that collects information after many web searches (web crawls!) and displays it in an infobox next to search results. This is a powerful feature of indexing which organizes the most relevant content and features it in a separate box to ease the process of searching.



Google uses RankBrain, a machine learning-based search algorithm for ranking webpages. It understands search queries and measures how much users are satisfied with the results by checking the user interaction with the displayed search results. The algorithm learns from this experience (just like a human!) and displays more user satisfactory and relevant results next time when a user gives a similar query. In essence, Google's Knowledge Graph and RankBrain usually work together to produce the best results for the user. The amazing part of Google's search algorithm is that it performs these actions within fractions of seconds.



Search Engine Optimization (SEO)

Before summarizing what we learned in this article, let's take a brief look at why understanding search engine mechanism is so important, and discuss what Search Engine Optimization (SEO) is.

Search engine optimization is used by business owners and website owners so that the search engine crawls their site more frequently, and so that it appears at the top of search engine results. Strategies are used to optimize, update and produce relevant content on the website so that it is displayed more frequently on the Search Engine Result Page (SERP).

In general, if one wants their website to get a higher rank, they should focus on the features that search engines use to rank pages that we discussed in earlier sections. Improving user experience, using more keywords, producing high-quality content along with proper layout are some of the important points to keep in mind when optimizing a webpage for search engines. These changes would allow the search engine crawlers to locate the site easily and give a higher rank as compared to other sites.

Summary

In this lesson, we learned about the working of search engines and also looked into some features of Google's search engine. SEO was also briefly introduced, which highlighted the importance and the need of understanding the search engine mechanism.