

Search engines

Table of contents

[Learning objectives](#)

[1. Different search engines](#)

[2. How do search engines work?](#)

[Advanced: RankBrain](#)

[Advanced: Important Google updates](#)

[Key takeaways](#)



Learning objectives

In this lesson, you'll learn:

- about the most influential search engines;
 - how search engines find your site and determine its ranking on the search results page;
 - about the most important Google updates and their implications;
-

1. Different search engines

There's a world of search engines at people's disposal by the click of a mouse. Let's take a look at some different options. The most popular search engine is **Google**, which has a reported market share of 90% globally. That means that it's undeniably the most popular search engine at this point. But, what else is there? Well, what about **Bing**? It's the second largest search engine in North America. About 6% of people in the US and Canada use Bing as their main search engine. Also, 4% of US searchers use **Yahoo**. This

means that it's worth researching what search engine your target audience uses if you live in the States.

But, there's more. There are regions where other search engines are more dominant. In Russia, for example, **Yandex** leads the pack with 45% of Russian market share. Also, the market leader in China is **Baidu**, which accounts for 65% of the market share over there. There's also privacy-oriented search engines, such as **Duckduckgo** and **Startpage**. They hold less than 1% of market share. Yet, some people will only use *these* because of their views on personal data. So it's good to know that they exist.

Lastly, we'd like to mention **Ecosia**. This search engine spends its profits [planting trees](#) to benefit the environment and local economy. Last time we checked, it helped plant over 77 million trees. Isn't that awesome?

Of course, stats change over time. Unsure about the current user numbers? Use [StatCounter](#) to check today's situation.

In our courses, we focus on **Google** because it's the largest search engine, the most well-known, and it has the most users globally. At Yoast, when we think of Google, we think of its mission, which is:

"to organize the world's information and make it universally accessible and useful."

Google's mission gives us information about what they think is important. But how exactly do search engines work?

2. How do search engines work?

Search engines follow **links**. Links connect pages and documents, much like roads connect villages and cities. By following these links, search engines collect data to show to their users. But how do they do it?

Well, a search engine like Google consists of:

1. spiders, also called bots or crawlers
2. an index
3. an algorithm

We'll explore these terms in further detail below.

Spiders, crawlers or bots

A **spider** follows the links on the web. A spider is also called a crawler, robot or bot. It goes around the internet 24/7. You can think of a spider as a little, curious explorer. It collects content such as text, images, videos, news items, and anything that'll be interesting for a search engine to show.

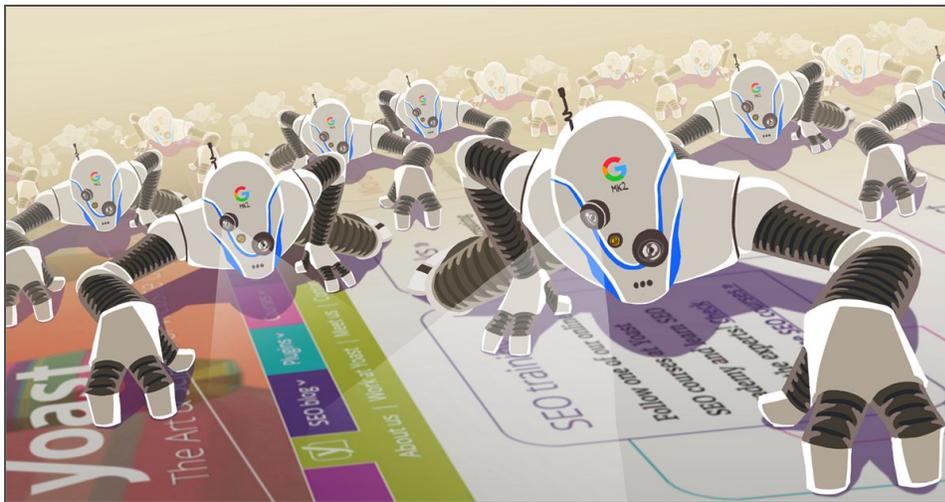


Image 1: Spiders

Once it comes to a website, it saves the HTML version of a page in a gigantic database, called the **index**. This index is updated every time the spider comes around your website and finds a new or revised page on it. Depending on how important Google deems your site and the number of changes you make on your website, the spider comes around more or less often.

The search engine algorithm

After indexing your website, search engines can show your website in the search results. But how do they decide what goes where? Well, search engines have an **algorithm** that does this for them. The algorithm takes the data from the index and calculates which site the user would most likely want to view based on their query.

It is very difficult to discover what factors into a search engine algorithm like Google's. Nobody knows exactly which factors decide how Google determines search results. Moreover, factors and their importance change very often. Testing and experimenting gives us at Yoast a good feel for the important factors and the changes in these factors.

How do search engines find your site?

For a search engine to know of the existence of your website, there first has to be a link from another site – one they already know – to your site.

Following that link will lead to the first crawler-session and the first save in the index. In the old days, you could submit your website to a search engine. Today, that isn't possible anymore. Search engines nowadays follow all the links on the web.

The value of links for search engines

It's very important to have a basic understanding of how Google and most other search engines use **links**: they use the number of links pointing to a page to determine how important that page is. Both internal links (from the own website), as well as external links (from other websites), can help a page rank high.

Some links are more important than others: links from websites that have a lot of external links themselves are generally more important than links from small websites with few external links.

Crawlability

Depending on how you maintain your website, it can be easy or difficult for Google to crawl your website. If you have good **crawlability**, Google will be able to index your site without any problems. There are a few ways in which a crawler can be blocked from your website. If the website or a page on your website is blocked, you're telling Google's spider: "do not come here, this area is forbidden". These pages won't turn up in the search results in most of these cases.

There are several ways you can prevent Google from accessing certain pages. If you want to know more about these methods, check out our Technical SEO course, part of the [Yoast SEO academy Premium subscription](#).



Advanced: RankBrain

RankBrain is a Google algorithm, but a very advanced one at that. It is a machine learning system that helps Google better decipher the meaning behind the terms people search for. It serves the best-matching search results related to those queries. When RankBrain was first announced, Google called it the third most important ranking factor. Presumably, RankBrain can somehow summarize what a page is about, evaluate the relevance of search results, and teach itself to get even better at it with time. The common understanding is that RankBrain, in part, relies on the traditional SEO factors (links, content, keywords, etc.), but also looks at other factors that are specific to the search term. Then, it identifies the most relevant pages in the index and arranges the respective results in SERPs (search engine result pages).



Advanced: Important Google updates

Throughout the years, Google introduced several major updates. We'll discuss the most important ones of the last seven years and the implications of these updates.

Before we start, let's talk about how easy it is to get caught up in all the changes. But, keeping track of every little update is energy that you could be putting into creating great copy or taking measures to reduce site loading time, for example. Google's priority is serving its users, so work to be the best result.

Panda (2011)

In 2011, Google released its first **Panda update**. This Panda update tried to diminish those websites which were purely created to rank in the search engines. Panda mostly focused on **on-page factors**. In other words, it determined whether the site offered information about the search term visitors used. Two types of sites were hit especially hard by the Panda update:

1. affiliate sites (sites that mainly exist to link to other pages)
2. sites with very thin content

Google has periodically re-run the Panda algorithm since its first release.

Penguin (2012)

A year later, Google rolled out the first **Penguin update**. Penguin particularly looked at the **links** websites got from other sites. It judged whether the sites linking to your website like and admire your products or content. If the links were artificial, Google no longer assigned link value. In the past, lots of people tried to boost their ranking by buying links. Penguin tried to diminish the effect of buying, exchanging or otherwise artificially creating links. Websites with a lot of these artificial links got hit hard by this update. They lost their place in Google's ranking. This update has also run several times since its first inception and is now even said to be run continuously.



Image 2: Panda and Penguin updates

Hummingbird (2013)

In August 2013, Google released **Hummingbird**. In this update, Google laid down the groundwork for **voice search**. Hummingbird pays more attention to each individual word in a query, ensuring that the whole search phrase is taken into account, rather than just particular words. This should lead to results matching the entire query better. The results were not immediately clear, but over time, Google started showing more answer boxes in the search results (see Image 3), that gave the answer directly instead of enticing people to click on to a web page. Voice search has become more and more important as more devices (Siri, Google Home, Alexa) have started using voice search.

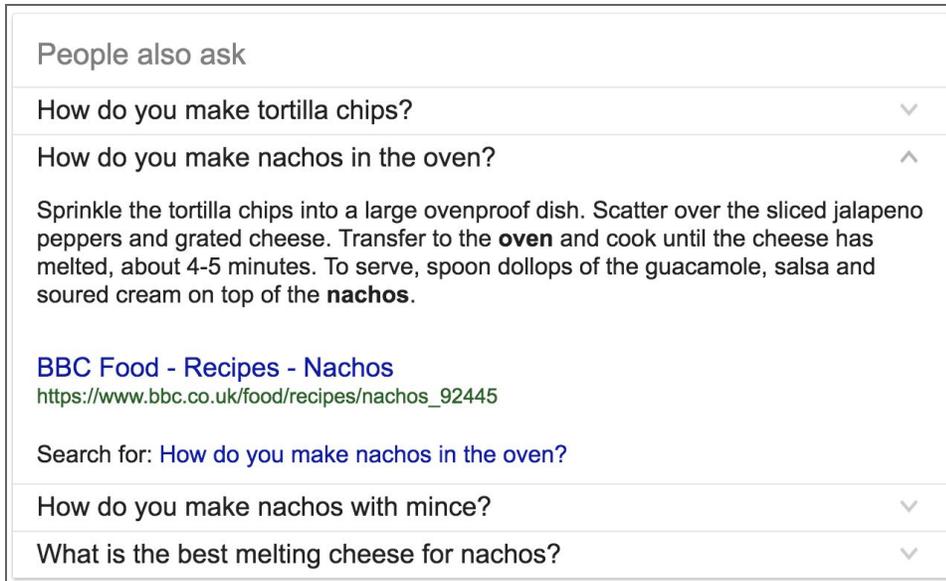


Image 3: Answer box in Google's search results

Mobilegeddon (2015)

In 2015, Google introduced the mobile update, dubbed **Mobilegeddon** by the industry. As more and more people use Google on mobile devices, Google used this update to boost sites that have **mobile-friendly pages** in its mobile search results. Around the same time, Google showed that mobile devices accounted for over 50% of all search queries.

Possum (2016)

In September 2016, the **Possum update** applied several recent changes to Google's **local ranking filter**. After Possum, Google has shown more varied results depending on the physical location of the searcher (the closer you are to a business physically, the more likely it is that you'll see it among local results) and the phrasing of the query (even very similar variations now produce different results).



Image 4: Possum update

Mobile-first index (2018)

In 2018, Google switched to its **mobile-first index**. This means Google will determine rankings based on the quality of the **mobile version of the site** instead of the desktop version. As more and more searches happen on mobile, Google wants its index and results to represent the majority of their users, who are mobile searchers. A new Googlebot will crawl your mobile site and determine if its performance, content, and user experience are up to scratch. If so, you get a better ranking. If it's lacking, other sites will rank higher, and you could lose out.

Medic, or Query Intent Update (2018)

In the same year, Google released a new update. It was called **Medic** by reporters, for it seemed to target medical websites at first. However, later, it was discovered that the update affected organic results in various industries. It seemed that this update changed the algorithm to benefit sites that matched their top content with **how a user would ask their question**.

Let's look at an example: "What is the best way to treat a burn?". In this case, Google's algorithm can now attempt to deduce the intent of the query. The user wants to compare remedies for a burn to know the best one, so Google wants to provide the site that gives the absolute best quality answer to this

question. It's a different search intent than if you were to search for "treat burn". In that case, Google suggests a how-to guide for burn treatment.

This change eventually resulted in people referring to the update as the **Query Intent Update**.

Continuous Updates (2018)

Since 2018, Google has begun updating continuously, resulting in multiple small adjustments each day. Also, Google continues to release bigger updates, also known as **core updates**, every few months. This means that it becomes increasingly hard to predict the ranking algorithm. So, we advise you to try and focus on making your site useful for people. If you align your goals with Google's mission, you have a better chance at ranking.

BERT (2019)

Last, but not least, Google has released a **natural language programming model** called BERT. BERT stands for **Bidirectional Encoder Representations from Transformers**. Those are a lot of difficult words to say that Google is now better at reading and understanding texts. Before, Google could only interpret text in one direction. Whereas now, it can associate words in both directions simultaneously. This makes it much easier for Google to draw accurate context from texts. In turn, this results in better search results. By using BERT, Google can link the appropriate context of different sites to the right search query.

The BERT algorithm is also used by other search engines, such as Bing.

Key takeaways

- Currently, the most influential search engine is Google. But, Bing, Yandex and Baidu are also regionally relevant. You can use [StatCounter](#) to check how user traffic is split between search engines globally.

- Search engines use crawlers to **find your site** by following links. If they find your site, it will be saved in the index. After indexing your website, search engines can show your website in the search results. The position in the search results is decided by a secret algorithm.
 - There have been several important Google updates that influenced the way we practise SEO.
-