

The Internet: How Search Works

[Code.org](#)

13 Jun 2017

5:12 min

Transcript

Hi, my name's John. I lead the search and machine learning teams at Google. I think it's amazingly inspiring that people all over the world turn to search engines to ask trivial questions and incredibly important questions. So it's a huge responsibility to give them the best answers that we can.

Hi, my name's Akshaya and I work on the Bing search team. There are many times where we will start looking into artificial intelligence and machine learning, but we have to address how are the users going to use this, because at the end of the day, we want to make an impact to society.

Let's ask a simple question. How long does it take to travel to Mars? Where did these results come from and why was this listed before the other one?

Okay, let's dive in and see how the search engine turned your request into a result.

The first thing you need to know is when you do a search, the search engine isn't actually going out to the World Wide Web to run your search in real time.

And that's because there's over a billion websites on the internet and hundreds more are being created every single minute. So if the search engine had to look through every single site to find the one you wanted, it would just take forever.

So to make your search faster, search engines are constantly scanning the web in advance to record the information that might help with your search later. That way, when you search about travel to Mars, the search engine already has what it needs to give you an answer in real time.

Here's how it works. The internet is a web of pages, connected to each other by hyperlinks. Search engines are constantly running a program called a Spider that cross

through these web pages to collect information about them. Each time it finds a hyperlink, it follows it until it has visited every page it can find on the entire internet. For each page the spider visits, it records any information it might need for a search by adding it to a special database called a search index.

Now, let's go back to that search from earlier and see if we can figure out how the search engine came up with the results. When you ask how long does it take to travel to Mars, the search engine looks in each of those words in the search index to immediately get a list of all the pages on the internet containing those words.

But just looking for these search terms could return millions of pages, so the search engine needs to be able to determine the best matches to show you first. This is where it gets tricky because the search engine may need to guess what you're looking for.

Each search engine uses its own algorithm to rank the pages based on what it thinks you want. The search engine's ranking algorithm might check if your search term shows up in the page title, it might check if all of the words show up next to each other, or any number of other calculations that help it better determine which pages you'll want to see and which you won't.

Google invented the most famous algorithm for choosing the most relevant results for a search by taking into account how many other Web pages linked to a given page. The idea is that if lots of websites think that a web page is interesting, then it's probably the one you're looking for. This algorithm is called page rank, not because it ranks web pages, but because it was named after its inventor, Larry Page, who's one of the founders of Google.

Because a website often makes money when you visit it, spammers are constantly trying to find ways to game the search algorithm so that their pages are listed higher in the results. Search engines regularly update their algorithms to prevent fake or untrustworthy sites from reaching the top.

Ultimately, it's up to you to keep an eye out for these pages that are untrustworthy by looking at the web address and making sure it's a reliable source. Search programs are always evolving to improve the algorithms so they return better results, faster results than their competitors.

Attribution: [The Internet: How Search Works \(video\)](#) by [Code.org](#)

Today's search engines even use information that you haven't explicitly provided to help you narrow down your search. So, for example, if you did a search for dog parks, many search engines would give you results for all the dog parks nearby, even though you didn't type in your location.

Modern search engines also understand more than just the words on a page, but what they actually mean in order to find the best one that matches what you're looking for. For example, if you search for fast pitcher, it will know you're looking for an athlete. But if you search for large pitcher, it will find you options for your kitchen. To understand the words better, we use something called machine learning, a type of artificial intelligence. It enables search algorithms to search out not just individual letters or words in the page, but understand the underlying meaning of the words.

The internet is growing exponentially, but if the teams that design search engines do our jobs right, the information you want should always be just a few keystrokes away.