

- » Discovering where people search
- » Understanding the difference between search sites and search systems
- » Distilling thousands of search sites down to three search systems
- » Understanding how search engines work
- » Gathering tools and basic knowledge

Chapter **1**

Surveying the Search Engine Landscape

You've got a problem. You want people to visit your website; that's the purpose, after all — to bring people to your site to buy your product, or find out about your service, or hear about the cause you support, or for whatever other purpose you've built the site. So you've decided you need to get traffic from the search engines — not an unreasonable conclusion, as you find out in this chapter.

So where do you start? You know you want to have your site appear in Google, of course. . . . but as big as Google is, it isn't everything. A lot of searches are carried out at sites *other than* Google. But when you start to consider other search options, the field starts to get crowded. There's AOL.com, Yahoo.com, and Bing.com, of course. But there's more; what about DuckDuckGo (a search site focused on privacy), DogPile.com, Ask.com, Baidu.com, Yandex.com, StartPage.com, and SwissCows.com?

And don't forget the nontraditional "search engines." Many searches are carried out at Amazon, Craigslist, eBay, and other shopping-related sites. Then there's sites such as YouTube, Twitter, and Facebook.

So where do you direct your attention? Well, I've got some good news. While you definitely need to consider more than just Google, the overall picture of search engine "targets" can be simplified. The point of this chapter is to take a complicated landscape of scores, maybe hundreds, of search sites and whittle it down into the small group of search engines that really matter. (Search sites? Search systems? Don't worry; I explain the distinction.)

Investigating Search Engines and Directories

The term *search engine* has become the predominant term for *search system* or *search site*, but you need to understand the different types of search, um, thingies that you're going to run across.

Although out on the Interwebs you will hear the term *search engine* a lot, perhaps almost exclusively, I like to sometimes use the term *search site*. Why? Because many search sites don't have their own search engines; rather, they partner with a search engine to provide their site visitors with search results.

Take, for instance, AOL.com (www.aol.com). One may be forgiven for thinking that AOL.com is a search engine; after all, it has a big search box right at the top, and if you enter a phrase and press Enter, or click a colored SEARCH button, you get search results.

INDEX ENVY

A few years ago, Yahoo! and Google used to compete to see who had the largest index; Google used to even publish the number of indexed pages on its home page; at one point the statement under the search box said that Google had indexed 15 billion pages.

Oh, the good old days . . . how things have changed. Now Yahoo! no longer has its own index (it gets search results from Bing), and forget billions of pages; now Google has found *trillions* of pages! In 2015, Google reported that it had discovered 60 trillion pages, though not all were indexed; today Google's How Search Works page (www.google.com/search/howsearchworks) states that the index itself contains "hundreds of billions of pages," and contains about 100 million gigabytes of data. (It's been saying that for a couple of years at least, so we're getting into fuzzy number territory. The bottom line? The index is *yuge!*)

However, AOL doesn't own a search engine, despite the fact that you can search at the AOL site. (Indeed, *many* people *do* search at AOL, around a couple of hundred million times a month). Rather, AOL gets its search results from the Bing search engine. Another example is EarthLink.net; this site (owned by an Internet Service Provider that used to be one of the top companies back in the 1990s) has a search box, but the search results come from Google. Hence my desire to differentiate between search *sites* (places where you can search) and search *engines* (the systems that actually do all the work). It's an important distinction.

Search sites, indexes, and engines

Let me quickly give you a few simple definitions:

- » **Search site:** A website where you can search for information on the web.
- » **Search engine:** A system that collects pages from the web, saves them in a massive database, indexes the information, and provides a mechanism for people to search through the data.
- » **Search index:** The index containing all the information that the engine collected and searches.
- » **Search directory:** A system that contains some basic information about websites, rather than about collected and indexed web *pages*.



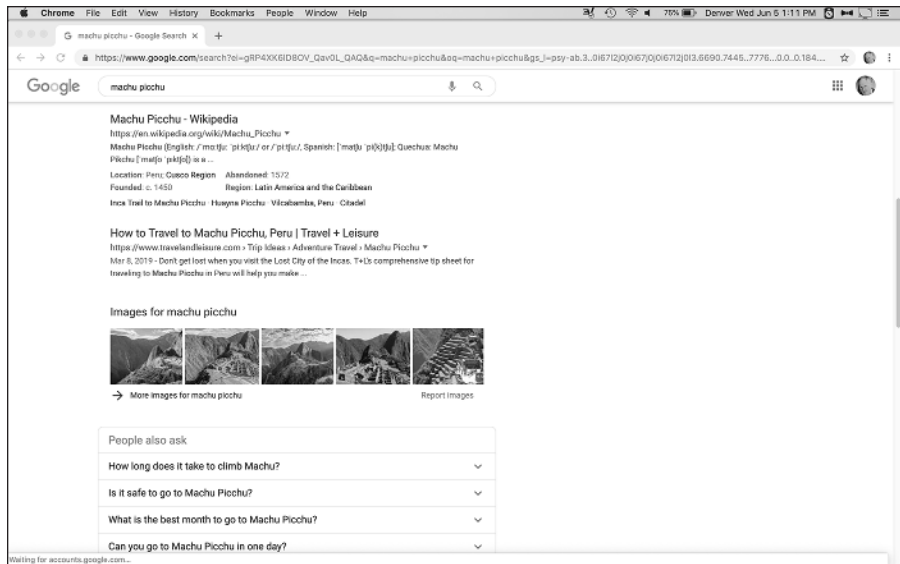
TECHNICAL
STUFF

Large search-index companies own thousands of computers that use software known as *spiders*, *searchbots*, or *robots* (or just plain *bots*) to grab web pages and read the information stored in them. These systems use complex *algorithms* — calculations based on complicated formulae — to index that information and rank it in search results when people search. Google, shown in Figure 1-1, is the world's most popular search site.

Search directories

Before there were search engines, there were search directories. A *directory* is a categorized collection of information about websites. Rather than containing information *from* web *pages*, it contains information *about* websites. In fact, before Google was even a twinkle in its fathers' eyes, Yahoo! directory was America's dominant search site; "The Google of the 1990s," as I've seen it described.

FIGURE 1-1:
Google, the world's most popular search engine, produced these results.



Directories are not created using spiders or bots to download and index pages on the websites in the directory; rather, for each website, the directory contains information, such as a title, description, and category, submitted by the site owner. The most significant search directories in recent years were owned by Yahoo! (<http://dir.yahoo.com>) and the Open Directory Project, affectionately known as DMOZ (pronounced “dee-moz”) due to its original name — Directory Mozilla — and its domain name, www.dmoz.org; see Figure 1-2; the Open Directory Project was actually a volunteer-managed directory owned by AOL. (You can see an archived version at <https://dmoz-odp.org/> if you’re interested.) Google used to have a directory, based on DMOZ data, at <http://dir.google.com>, but that’s long gone.

These directories had staff members who examined all the sites in the directory to make sure they were placed into the correct categories and met certain quality criteria; Yahoo! charged \$299 a year for the privilege of being listed in their directory.

However, search directories are simply nowhere near as important today as in the past. Yahoo! directory has gone, Google stopped using DMOZ data, and not long after that DMOZ itself closed its doors.

These directories became irrelevant to average users; most users didn’t even know they existed . . . and now they don’t.

However, directories *may* still be useful to your SEO efforts. There are still thousands of small, specialized directories, focusing on particular industries, hobbies, jobs, sports, cities, and so on, and these directories can be an important way to get traffic to your site. Chapter 16 addresses this topic.



FIGURE 1-2:
The Open Directory Project, back in its heyday.

Spidered directories

I wasn't sure what to call these things, so I made up a name: *spidered directories*. As I describe in the preceding section, directories don't contain a full index of the content of the sites pages; rather, they contain a little information about the site itself. In most cases the person who enters the site into the index provides this information, or sometimes a staff member does so. But a number of small search sites actually use *spiders* (searchbots) to grab a little background information about each site, and even pages within the site, such as titles, descriptions, and keywords. In some cases, this information comes from the meta tags pulled off the pages in the index. (I tell you about meta tags in Chapter 3.) So these are a form of directory, but they are generally created programmatically rather than by site owners requesting inclusion. (Yahoo! Directory and DMOZ were "hand built" using data submitted by site owners.) A number of the smaller systems discussed in Chapter 16 are of this type.

Pay-per-click systems

Many search sites provide *pay-per-click* (PPC) listings. When you search at Google, for instance, you'll see results that come out of Google's main index, but also small text ads. Advertisers place these small ads into the PPC system, and when

users perform their searches the results contain some of these sponsored listings, typically above and below the free listings. We won't be covering these ads in this book, as it's a very different subject from SEO.



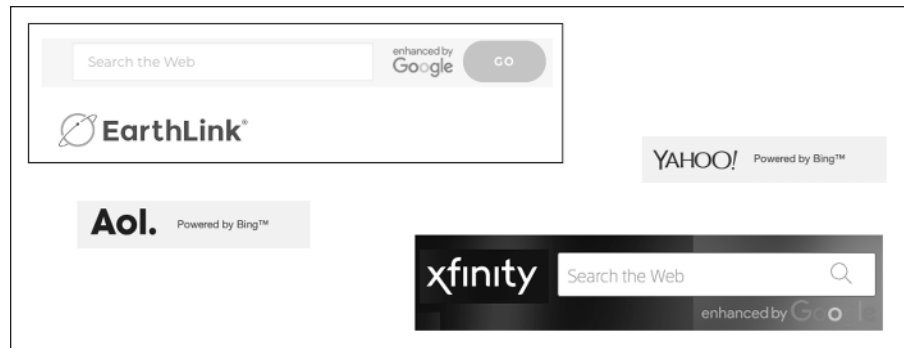
TIP

Keeping the terms straight

Here are a few additional terms that you'll see scattered throughout the book:

- » **Search term:** This is the word, or words, that someone types into a search engine when looking for information. Also frequently known as the *search query*.
- » **Keywords:** The flip side of *search term*. The terms you put into your web pages (I explain where in this book) in order to rank well for the matching search term. It's just a matter of terminology, and in fact you'll often hear people talk about entering keywords into the search box. When someone types *blue widget* into a search box, that's a search term in use; but you need the keyword *blue widget* in your web pages so that when someone uses that search term, your pages have a chance of being included in the search results.
- » **Search results:** Results are the information (the results of your search term) returned to you when you go to a search site and search for something. As explained earlier in this chapter, in many cases the search results you see don't come from the search site you're using, but rather from some other search index (see Figure 1-3).
- » **SERPs:** I don't use the term much, but you'll hear others in the business talking about the *serps*. It simply means *search engine results page*, the page that appears after you search and contains the search results.
- » **Natural or organic search results:** A link to a web page can appear on a search results page two ways: The search engine may place it on the page because the site owner paid for it to be there (pay-per-click ads), or it may pull the page from its index because it thinks the page matches the search term well. These free placements are often known as *natural search results*; you'll also hear the term *organic search results* and sometimes even *algorithmic search results*.
- » **Search engine optimization (SEO):** Search engine optimization (also known as *SEO*) refers to "optimizing" websites and web pages to rank well in the search engines — the subject of this book, of course.

FIGURE 1-3:
Look carefully,
and you'll see
that many search
sites get their
search results
from other search
systems.



Why bother with search engines?

Why bother using search engines for your marketing? Because search engines represent the single most important source of new website visitors.

You may have heard that most website visits begin at a search engine. Well, this isn't true, though many people continue to use these outdated statistics because they sound good — “80 percent of all website visitors reach the site through a search engine,” for instance. However, way back in 2003, that claim was finally put to rest. The number of search-originated site visits dropped below the 50 percent mark. Most website visitors reach their destinations by either typing a *URL* — a web address — into their browsers and going there directly or by clicking a link on another site that takes them there. Most visitors *don't* reach their destinations by starting at the search engines.

However, search engines are still *extremely important* for a number of reasons:

- » The proportion of visits originating at search engines is still significant. Sure, it's not 80 percent, but with billions of searches each month, it's still a lot of traffic. Way back in 2016 it was reported that Google alone gets *at least* 63,000 searches *a second* . . . that's over *5 billion* a day! More than *2 trillion* a year.
- » Many billions more searches are carried out in other search sites, such as major search sites like Bing.com and Yahoo.com, map sites (MapQuest), video sites (YouTube), retail sites (Amazon, eBay, Craigslist), and so on. It's likely that more than hundreds of billions searches are performed in the United States each month, dozens of searches every day for every man, woman, child, and baby in the United States.
- » Of the visits that don't originate at a search engine, a large proportion are revisits — people who know exactly where they want to go. This isn't new business; it's repeat business. Most *new* visits come through the search engines — that is, search engines are the single most important source of *new* visitors to websites.

- » It's also been well established for a number of years that most people researching a purchase begin their research at the search engines. (Except for those who don't. As I discuss in Chapter 17, many, perhaps most, product searches actually begin in sites such as Amazon, eBay, and Craigslist. But then, I think it's important to understand that these sites *are* search engines; they are, in effect, product-search engines.)
- » Search engines represent an inexpensive way to reach people. Generally, you get more bang for your buck going after free search-engine traffic than almost any other form of advertising or marketing.

Here's an example. One client of mine, selling construction equipment to the tune of \$10,000 a month, rebuilt his site and began a combined natural-search and paid-search campaign, boosting sales to around \$500,000 a month in less than two years. It's hard to imagine how he could have grown his company, with relatively little investment, so quickly without the search engines!

Where Do People Search?

You can search for websites at many places. Literally thousands of sites, in fact, provide the ability to search the web.

However, *most* searches are carried out at a small number of search sites. How do the world's most popular search sites rank? That depends on how you measure popularity:

- » Percentage of site visitors (*audience reach*)
- » Total number of visitors
- » Total number of searches carried out at a site
- » Total number of hours visitors spend searching at the site

Each measurement provides a slightly different ranking. Although all provide a similar picture with the same sites generally appearing on the list, some search sites are in slightly different positions.

The following list shows the United States' recent top general search sites (according to Alphametics and Statcounter. . .but, note that all these kinds of statistics are pretty fuzzy numbers, it's next to impossible to get a truly accurate number):

Google: 87.3 percent

Bing: 6.9 percent

Yahoo!: 4.6 percent

Other: 1.2 percent

Note that Yahoo! gets its results from Bing, thanks to a Yahoo!/Microsoft partnership — known as the *Yahoo! and Microsoft Search Alliance* — that was implemented in August 2010. (Look for the little *Powered by Bing* notice at the bottom of Yahoo! search pages.) However, it was rumored several years ago that Yahoo! wanted out of the agreement — so it could go back to using Google search results! — but couldn't figure out how to break the ten-year contract with Microsoft. By the time you read this book, the contract may have expired already, and Yahoo! may have switched . . . who knows?

What's in that “Other” category? Well, it barely matters, it's so small. But, just for the record, it includes DuckDuckGo.com (maybe one search out of every 200 goes through this site), Ask.com (which, long ago, used to be AskJeeves.com — remember? — and at one point had a few percent of the search market), and many others.

Note that these numbers are for the United States of America. So, a few things to consider when considering SEO in other locations:

- » Google is the world's dominant search engine; in much of the rest of the world, it's still the top, and worldwide more searches run through Google than any other search engine.
- » In China, though, the top search engine is Baidu (www.Baidu.com), distantly followed by Shenma.
- » In Russia, the dominant search engine is Yandex, closely followed by Google. Yandex is a significant player in a number other countries, such as Turkey and the Commonwealth of Independent States (the old Soviet Union).
- » In most of the world, though, Google's dominance rivals that of Google's dominance in the U.S., often in the region of 90 percent (and often considerably more) of searches carried out every day.
- » Outside Russia and China, in general Bing is the #2 search engine, just as it is in the United States (a huge share for Google followed by a much smaller share for Bing).



REMEMBER

So where does that leave us? In the United States, there are really only two search engines of any Remember, in my small list earlier in this section, we had four things: Google and Bing, then Yahoo! (but Yahoo! gets its results from Bing), and then “Other,” which is tiny and in any case, many sites within the Other group get their results from Google and Bing. So all we really care about — in the United States — are Google and Bing!

That's not so bad, is it? When I first wrote *SEO For Dummies* back in 2003 (well, it was *Search Engine Optimization For Dummies* in those days), there were thousands of search engines, search sites, and directories. Most are long gone. DMOZ has gone, and all the hundreds of directories that worked with DMOZ data. Yahoo! directory has gone, Yahoo! uses Bing data, Altavista has gone (and at one point Altavista was one of the world's top search engines), HotBot has gone, Mamma has gone. AllTheWeb just redirects to Yahoo! And many of them that are still around use other people's data; Infospace, Dogpile, and Webcrawler, for instance, get their data from Google and Bing.

It's all the same!

But here's the thing. Not only are there only really two search engines you need to care about now, they both work in a *very* similar manner! While it used to be popular to talk about "Google SEO," and "Bing SEO," they are really pretty much the same thing. Optimizing your site for Google is the same as optimizing your site for Bing. Both search engines are interested in well-written page content, keywords placed into pages, keywords in links pointing back to your site, keywords in URLs, and so on. Optimize for one, and you are optimizing for the other.

However, there are other sites to consider. In some cases, you need to check out specialty directories and indexes related to the industry in which your website operates or submit your site to web directories in order to build links back to your site. (You learn about using directories in Chapter 16, and the value of links pointing to your site in Chapters 18 – 20.) In addition, in Chapter 17, you find out about the product search sites — hugely important for those of you selling products. And in Chapter 22, I tell you about the video sites — YouTube, for instance, is one of the world's most important search engines, after Google and Bing.



REMEMBER

Get your site indexed by both Google and Bing, and you're in front of probably around 99 percent of all searchers (in the general-search sites, that is). Well, *perhaps* you're in front of them. You have a *chance* of being in front of them, anyway, if your site ranks highly (which is what this book is all about).

Search Engine Magic

Go to Google and search for the term *personal injury lawyer*. Then look at the blue bar below the Google logo, and you see something like this:

About 267,000,000 results (0.81 seconds)

This means Google has found over 267 million pages that it believes match these three words in some way. Yet, somehow, Google has managed to rank the pages. It's decided that one particular page should appear first, and then another, and then another, and so on. (By the way, this has to be one of the wonders of the modern world: Search engines have tens of thousands of computers, evaluating a trillion pages or more, in a fraction of a second.)

How do they do it?

How on earth does Google do it? How does it evaluate and compare pages? How do other search engines do the same? Well, I don't know *exactly*. Search engines don't want you to know how they work (or it would be too easy to create pages that exactly match the criteria of the search system for any given search term, "giving them what they want to see"). But I can explain the general concept.

When Google searches for your search term, it begins by looking for pages containing the exact phrase. Then it starts looking for pages containing the words close together, and for synonyms; search for *dog* and Google knows you may be interested in pages with the word *canine*, for instance. (One Google source claims that synonyms come into play in around 70 percent of all searches.) Then it looks for pages that have the words scattered around. This isn't necessarily the order in which a search engine shows you pages; in some cases, pages with words close together (but not the exact phrase) appear higher than pages with the exact phrase, for instance. That's because search engines evaluate pages according to a variety of criteria.

Search engines look at many factors. They look for the words throughout the page, both in the visible page and in the nonvisible portions of the HTML source code for the page. Each time they find the words, they are *weighted* in some way. A word in one position is worth more than a word in another position. A word formatted in one way is worth more than a word formatted in another. (You can read more about this in Chapter 7.) There's more, though. Search engines also look at links pointing to pages and use those links to evaluate the referenced pages: How many links are there? How many are from popular sites? What words are in the link text? You read more about this in Chapters 18 through 20.

Stepping into the programmers' shoes

There's a lot of conflicting information out there about SEO. Some of it's good, some of it's not so good, and some of it's downright nonsense. When evaluating a claim about what search engines do, I sometimes find it useful to step into the shoes of the people building the search engines; I try to think about what would make sense from the perspective of the programmers who write the code that evaluates all these pages.

Consider this: Say, you search for *personal injury lawyer*, and the search engine finds one page with the term in the page’s title (between the <TITLE> and </TITLE> tags, which you can read more about in Chapters 3 and 7), and another page with the term somewhere deep in the page text. Which do you think is likely to match the search term better? If the text is in the title, doesn’t that indicate that page is likely to be related in some way to the term? If the text is deep in the body of the page, couldn’t it mean that the page isn’t directly related to the term, but that it’s related to it in some incidental or peripheral manner?



TIP

Considering SEO from this point of view makes it easier to understand how search engines try to evaluate and compare pages. If the keywords are in the links that point to the page, the page is likely to be relevant to those keywords; if the keywords are in headings on the page, that must be significant; if the keywords appear frequently throughout the page, rather than just once, that must mean something. Suddenly, it all makes sense.

By the way, in Chapter 11, I discuss things that search engines don’t like. You may hear elsewhere all sorts of warnings that may or may not be correct. Here’s an example: I’ve read that using a refresh meta tag to automatically push a visitor from one page to another will get your site penalized and may even get your site banned from the search engine. You’ve seen this situation: You land on a page on a website, and there’s a message saying something like, “We’ll forward you to page *x* in five seconds, or you can click [here](#).” The theory is that search engines don’t like this, and they may punish you for doing this.

Now, does this make any sense? Aren’t there good reasons to sometimes use such forwarding techniques? Yes, there are. So why would search engines punish you for doing it? They don’t. They probably won’t index the page that is forwarding a visitor — based on the quite reasonable theory that if the site doesn’t want the visitor to read the page, the search engine doesn’t need to index it — but you’re not going to get punished for using it.

Remember that the search engine programmers aren’t interested in punishing anyone; they’re just trying to make the best choices between billions of pages. Generally, search engines use their “algorithms” to determine how to rank a page, and they try to adjust the algorithms to make sure “tricks” are ignored. But they don’t want to punish anyone for doing something for which there might be a good reason, even if the technique could also be used as a trick.

What would the programmers do? I like to use this as my “plausibility filter” when I hear someone make some unusual or even outlandish claim about how search engines function.

Gathering Your Tools

You need several tools and skills to optimize and rank your website. I talk about a number of these in the appropriate chapters, but I want to cover a few basics before I move on. It goes without saying that you need:

- » Basic Internet knowledge
- » A computer connected to the Internet
- » A web browser
- » A website
- » One of these three things:
 - Good working knowledge of HTML
 - Access to a geek with a good working knowledge of HTML
 - A website creation tool that provides SEO functions that allow you to modify the site in the required manner

Certain changes need to be made to a website in order to optimize it properly; the Title tag needs to be changed, along with the Description meta tag, the headings need to use H1 tags, you need to be able to put keywords into the URL, and so on. This means that whoever does this work needs to understand what these things mean, and how to modify them. Or the tool you use to build your website has to provide a convenient way to allow you to change these elements. Some do, some don't.

Teaching HTML and how to upload pages to a website is beyond the scope of this book. If you're interested in finding out more, check out *HTML, XHTML, & CSS For Dummies*, by Ed Tittel and Jeff Noble, and *Creating Web Pages For Dummies*, 9th Edition, by Bud E. Smith (both published by John Wiley & Sons, Inc.).

All the major browsers (Chrome, Firefox, Internet Explorer, Safari, and Opera) have a bunch of SEO-related tools, which can often be useful. Some of these tools are provided by major SEO companies — companies that provide link-analysis, keyword-research, site-audit tools, and the like — such as Moz, Majestic, and AHrefs. It's worth spending a little time looking through your browser's add-on or extension library for tools such

- » **NoFollow tools:** Lots of tools indicate the presence of "nofollow" links (see Chapter 18).
- » **Whois tools:** These tools retrieve information about the domain of the site you're viewing — great for digging up info on competitors.

- » **Link analysis tools:** These things quickly provide information about links pointing to the site you are viewing — how many links, the value of the links, and so on.
- » **General purpose tools:** There's a bunch of tools that provide a wide range of information, often in a browser toolbar: estimates of the number of visitors coming to the site, information about social-media links to the site, siterank in Google and Bing, the number of pages from the site indexed in Google and Bing, information extracted from the site's HTML showing the title tag, description tag, and so on.

In fact, there are many SEO tool companies (see Chapter 27 for a bunch). I'm putting together a few special offers from some of these firms; for instance, you can get extended free trials from some. See www.SearchEngineBulletin.com for information.



TIP

GEEK OR NO GEEK

Many readers of this book are business people who don't plan to do the search engine work themselves (or, in some cases, realize that it's a lot of work and need to find someone with more time or technical skills to do the work). However, having read this book, they understand far more about search engines and are in a better position to find and direct someone else working on their site. As one reader-cum-client told me, "There's a lot of snake oil in this business," so his reading helped him understand the basics and ask the right questions of search engine optimization firms.