

# Chapter 1: Putting Search Engines in Context

---

## *In This Chapter*

- ✓ Identifying search engine users
- ✓ Discovering why people use search engines
- ✓ Pinpointing elements for getting high keyword rankings
- ✓ Defining relationships between search engines

The Internet offers a world of information, both good and bad. Almost anything a person could want is merely a few taps on the keyboard and a couple clicks of a mouse away. A good rule of thumb for the Internet is if you want to know about something or purchase something, there's probably already a Web site just for that. The catch is actually *finding* it. This is what brings you to this book. You have a Web site. You have hired what you hope is a crack team of designers and have unleashed your slick, shiny new site upon the Web, ready to start making money. However, there is a bit of a problem: Nobody knows that your site exists. How will people find your Web site?

The most common way that new visitors will find your site is through a search engine. A *search engine* is a Web application designed to hunt for specific keywords and group them according to relevance. It used to be, in the stone age of the 1990s, that most Web sites were found via directories or word-of-mouth. Somebody linked to your Web site from their Web site, or maybe somebody posted about it on one of their newsgroups, and people found their way to you. Search engines such as Google, Yahoo!, and Microsoft Live were created to cut out the middleman and bring your user to you with little hassle and fuss.

In this chapter, we show you how to find your audience by giving you the tools to differentiate between types of users, teaching you to sort out search engines, identifying the necessary elements for being prominent in those engines, and giving you an insider look at how all the search engines work together.

### *Identifying Search Engine Users*

Who is using search engines? Well, everyone. A significant amount of all Web traffic to Web sites comes from search engines. Unless you are a household name like eBay or Amazon, chances are people won't know where you are unless they turn to a search engine and hunt you down. In fact, even the big brands get most of their traffic from search engines. Search engines are the biggest driver of traffic on the Web and their influence only continues to grow.

But although search engines drive traffic to Web sites, you have to remember that your Web site is only one of several and a half trillion other Web sites out there. Chances are, if someone does a search, even for a product that you sell, your Web site won't automatically pop up in the first page of results. If you're lucky and the query is targeted enough, you might end up somewhere in the top 100 of the millions of results returned. That might be okay if you're only trying to share your vacation photos with your family, but if you need to sell a product, you need to appear higher in the results. In most cases, you want the number one spot on the first page because that's the site everyone looks at and that most people click.

In this section, you find out a bit more about the audience available to you and how to narrow down how to reach them.

#### *Figuring out how much people spend*

The fact of the matter is that people spend money on the Internet. It's frightfully easy: All you need is a credit card, a computer with an Internet connection, and something that you've been thinking about buying. E-commerce in the United States reached \$34.7 billion in the third quarter of 2007 alone. Some project that e-commerce could reach \$1 trillion a year by 2012. Combine that with the fact that most Americans spend an average of 24 minutes a day shopping online, not including the time they spend actually getting to the Web site (19 minutes), and you're looking at a viable means of moving your product. To put it simply, "There's gold in them thar hills!"

So, now you need to get people to your Web site. In real estate, the most important thing is location, location, location, and the same is true of the Internet. On the Web, however, instead of having a prime piece of property, you need a high listing on the *search engine results page (SERP)*. Your placement in these results is referred to as your *ranking*. You have a few options when it comes to achieving that. One, you can make your page the best it can be and hope that people will find you, or two, you can pay for one of the few advertising slots. More than \$12 billion was spent in 2007 on the North American search marketing industry alone. Eighty-eight percent of that was spent on *pay per click (PPC)* advertising, in which you pay to have search

engines display your ad. The other 12 percent goes to *search engine optimization (SEO)*. *SEO*, when properly done, helps you to design your Web site in such a way that when a user is doing a search, your pages appear on the first page of returned results, hopefully in the top spot. Your main focus in this book is finding out about *SEO*, but because they overlap somewhat, you pick up a bit of *PPC* knowledge here and there along the way.

### *Knowing your demographics*

In order to get the most bang for your *SEO* buck, you need to know the demographics for your Web visitors. You need to know who's looking for you, because you'll need to know where best to advertise. For example, if you're selling dog sweaters, it's probably not a great idea to advertise in biker bars. Sure, there might be a few Billy Bob Skullcrushers with a cute little Chihuahua in need of a cashmere shrug, but statistically, your ad would probably do much better in a beauty salon. The same goes for your Web site in a search engine. Gender, age, and income are just a few of the metrics that you'll want to track in terms of identifying your audience. Search engine users are pretty evenly split between male and female search engine users, with a few slight differences: 50.2 percent of Yahoo! users are female, whereas 53.6 percent of Google users are male. In terms of age brackets, the older set leans more towards using Ask.com, and the younger users wind up on Yahoo! and MSN.com most often. In fact, Ask.com is changing their focus in order to cater specifically to married women. Google reaps the highest number of users with an income of \$100,000 a year or more. Search engines even feed their results into other search engines, as you can in see our handy-dandy Search Engine Relationship Chart later in this chapter. Table 1-1 breaks down user demographics across the search engines for your reference.

**Table 1-1 User Demographics Across Major Search Engines**

	<i>Google</i>	<i>Yahoo! Search</i>	<i>MSN Search</i>
Female	46.58%	50.76%	54.26%
Male	53.42%	49.24%	45.74%
18-34	43.57%	48.23%	39.53%
35-54	42.85%	39.83%	44.49%
55+	13.57%	11.94%	15.99%
Under \$30K/year	20.00%	21.87%	21.01%
\$30K-100K/year	57.05%	57.69%	58.84%
Over \$100K/year	22.95%	20.44%	20.16%

*For the 12-week period ending May 15, 2004*

## 12 Identifying Search Engine Users

You need to know who your search engine visitors are because this demographic data helps you effectively target your market. This demographic distribution is often associated with search query *keywords*, the words that search engine visitors use to search for your products. For an in-depth look at choosing keywords, you can check out Book II, Chapter 2, but a brief summary is that *keywords* are what a search engine looks for when figuring out what sites to show in the SERP. Basically your keywords are the words you used in your *search query* — or what you typed into the little search window. If you are searching for something like information on customizing classic cars, for example, you would type [custom classic cars] into the search field. (When we discuss search queries through the book, we use square brackets to show the keywords. You wouldn't actually type the brackets into the search field.) Figure 1-1 displays a typical search engine results page for the query [custom classic cars].

**Figure 1-1:**  
Keywords  
in a search  
engine  
window:  
[custom  
classic  
cars].



The search engine goes to work combing its index for Web pages containing these specific keywords and returns to you with your results. That way, if you have a product that's geared towards a certain age bracket, or towards women more than men, you can tailor your keywords accordingly. It may seem inconsequential, but trust me, this is important if you want to be ranked well for targeted searches.

## Figuring Out Why People Use Search Engines

We've already established that a *lot* of people use search engines. But what are people looking for when they use them? Are they doing research for restoring their classic car? Do people use them to look for a place that sells parts for classic cars? Or are they just looking to kill time with video that shows custom cars racing? The answer is yes to all of the above. A search engine is there to scour the billions on billions of Web sites out there in order to get you where you need to go, whether it's doing research, going shopping, or just plain wasting time.

### 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. Someone doing a term paper on classic cars for their Automotive History 101 class would use it to find statistics on the number of cars sold in the United States, instructions for restoring and customizing old cars, and possibly communities of classic car fanatics out there. Companies would use it in order to find where their clients are, and who their competition is.

Search engines are naturally drawn to research-oriented sites and usually consider them more relevant than shopping-oriented sites, which is why, a lot of the time, the highest listing for the average query is a Wikipedia page. *Wikipedia* is an open-source online reference site that has a lot of searchable information, tightly cross-linked with millions of back links. *Open source* means that anyone can have access to the text and edit it. Wikipedia is practically guaranteed to have a high listing on the strength of its site architecture alone. (We go over site architecture in much more depth later on in Book IV.) Wikipedia is an open-source project, thus information should be taken with a grain of salt as there is no guarantee of accuracy. This brings us to an important lesson of search engines — they base “authority” on perceived expertise. Accuracy of information is not one of their criteria: Notability is.



In order to take advantage of research queries, you need to gear your site content toward things that would be of interest to a researcher. “How to” articles, product comparisons, reviews, and free information are all things that attract researchers to your site.

### Shopping

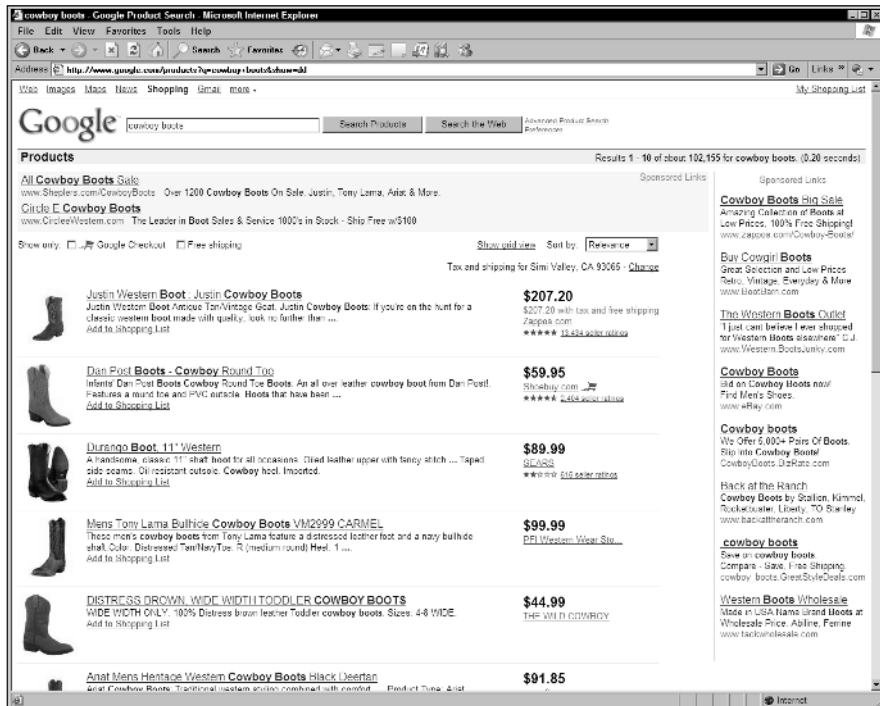
A smaller percentage of people, but still very many, use a search engine in order to shop. After the research cycle is over, search queries change to terms that reflect a buying mindset. Terms like “best price” and “free shipping” signal a searcher in need of a point of purchase. Optimizing a page to

meet the needs of that type of visitor results in higher *conversions* (actions taken by a user that meet a sales or business goal) for your site. As we mentioned, global search engines such as Google tend to reward research oriented sites, so your pages have to strike a balance between sales-oriented terms and research-oriented terms.

This is where specialized engines come into the picture. Although you can use a regular search engine to find what it is you're shopping for, some people find it more efficient to use a search engine geared directly towards buying products. Some Web sites out there are actually search engines just for shopping. Amazon, eBay, and Shopping.com are all examples of shopping-only engines. The mainstream engines have their own shopping products such as Google Product Search (formerly called Froogle) and Yahoo! Shopping, where you type in the search term for the particular item you are looking for and the engines return the actual item listed in the results instead of the Web site where the item is sold. For example, say you're buying a book on Amazon.com. You type the title into the search bar, and it returns a page of results. Now, you also have the option of either buying it directly from Amazon, or, if you're on a budget, you can click over to the used book section. Booksellers provide Amazon.com with a list of their used stock and Amazon handles all of the purchasing, shipping, and ordering info. The same is true of Yahoo! Shopping and Google Product Search. And like all things with the Internet, odds are that somebody, somewhere, has exactly what you're looking for. Figure 1-2 displays a results page from Google Product Search.

### ***Entertainment***

Research and shopping aren't the only reasons to visit a search engine. The Internet is a vast, addictive, reliable resource for consuming your entire afternoon, and 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. A child of the 80s might want to download an old-school version of the *Oregon Trail* video game onto her computer so she can recall the heady days of third grade. It's a quest made easy with a quick search on Google. Or if you want to find out what those wacky young Hollywood starlets are up to, you can turn to a search engine to bring you what you need.



**Figure 1-2:**  
A typical  
Google  
Product  
Search  
results  
page.

If you're looking for a video, odds are it's going to be something from YouTube, much like your research results are going to come up with a Wikipedia page. YouTube is another excellent example that achieves a high listing on results pages. They're an immensely popular video-sharing Web site where anyone with a camera and a working e-mail address can upload videos of themselves doing just about anything from talking about their day to shaving their cats. But the videos themselves have keyword-rich listings in order to be easily located, plus they have an option that also displays other videos. Many major companies have jumped on the YouTube bandwagon, creating channels for their companies (a YouTube *channel* is a specific account). Record companies use channels to promote bands, and production companies use them to unleash the official trailer for their upcoming movie.

### ***Discovering the Necessary Elements for Getting High Keyword Rankings***

If the mantra of real estate is location, location, location, and the very best location on the Web is on the search engines, the mantra of SEO should be keywords, keywords, keywords. Search engines use a process to categorize and grade keywords in order to bring you the Web pages you're looking for. The more relevant your keywords are to the user's query, the higher ranking your page has in a search engine's results. Keeping the keywords clear, precise, and simple helps the search engines do their job a whole lot faster. If you're selling something like customized classic cars, you should probably make sure your text includes keywords like classic cars, customized cars, customized classic Mustangs, and so forth, as well as clarifying words like antique, vintage, and restored. You can read more about how to choose your keywords in Book II.

In this section, you get a broad, brief overview on how you get a higher rank than the other guy who's selling macadamia nut butter. You need to know the basics, or you can't do targeted SEO.

#### ***The advantage of an SEO-compliant site***

Having an SEO-compliant Web site entails tailoring your Web site to have the highest SERP ranking for a keyword search. This includes optimizing your metadata and `Title` tag (for more on metadata, refer to Book IV, Chapter 3) so they are chock full (but not *too* full) of relevant keywords for your industry. Also, make sure that your Web page contains searchable text as opposed to lots of pretty Flash animations and images (search engines have limited ability to understand non-text content), that all of your images contain an *Alt attribute* (an alternative description of an image) with text that describes the content of the image, and that you have keywords embedded in your hyperlinks. You also need to be sure that all of your internal content as well as your links are siloed. You want to be sure to optimize every single one of these elements. Use this checklist to get yourself organized:

- ◆ `Title` tag
- ◆ `Meta description` tag
- ◆ `Meta keywords` tag
- ◆ `Heading` tag(s)
- ◆ Textual content
- ◆ `Alt` attributes on all images
- ◆ Strong/bold tags

- ◆ Fully-qualified links
- ◆ Site map
- ◆ Text navigation
- ◆ JavaScript/CSS externalized
- ◆ Robots text (.txt) file
- ◆ Web analytics
- ◆ Keyword research (technically a process — See Book II)
- ◆ Link development
- ◆ Image names
- ◆ Privacy statement
- ◆ Contact information
- ◆ Dedicated IP address

### ***Defining a clear subject theme***

Another way of getting a high keyword ranking is having a clear subject *theme*. If you're selling kits to customize classic cars, keeping your Web site streamlined and keeping all topics on the Web site relating exactly to classic car customization not only makes it easier for users to navigate your site and research or purchase what they need, but it also increases your chances of having a high page rank when those search engine spiders come by. The more similarly themed keywords you have on your pages, the better. It's the nature of a search engine to break up a site into subjects that add up to an overall theme for easy categorization, and the more obvious your site theme is, the higher your results will be.

It's kind of like going to an all-you-can-eat buffet and deciding you want to get a salad. You, the search engine, immediately go to the salad corner of the buffet because it's been clearly labeled, and from there, you can do your breakdowns. You want romaine lettuce, croutons, parmesan cheese, and Caesar dressing, so you go to where they keep the lettuce, the trimmings, and the dressings in the salad bar section. It's easy to find what you want if everything is grouped accordingly. But if the restaurant stuck the dressing over with the mashed potatoes, you'll have trouble finding it because salad dressing and mashed potatoes don't normally go together. Similarly, when you keep your Web site content organized with everything in its proper place, the search engine views your content with clarity, understanding what you're about — which in turn increases your page ranking. *Siloing* is a way of structuring your site and links in order to present a clear subject theme to the search engines. For more on this technique, refer to Book II, Chapter 4 as well as the entirety of Book VI.

### ***Focusing on consistency***

*Methodical consistent implementation* is the principle that, when you update your Web site, you should do it the same way every time. Your site should have a consistent look and feel over time without massive reorganizations at every update. In order for a search engine to maintain efficiency, you need to keep related content all placed in the same area. You also need to keep all of your updating processes consistent. That way, if something goes wrong during your next update, you can pinpoint what went wrong where without too much hassle since you update things the same way every time. It is confusing to customers to have things constantly changing around. Search engines and visitors to your Web site face the same challenge as a restaurant patron. Getting back to our salad bar analogy, the restaurant owner shouldn't scatter the salad dressings according to the whims of his salad bar designer, and randomly change things every time he gets in a new dressing or someone discontinues one of the old dressings.

### ***Building for the long term***

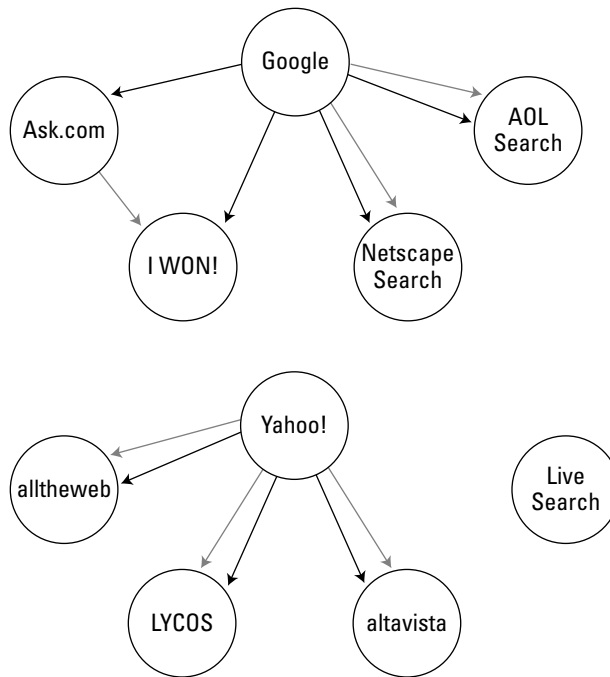
You need to consider your persistence for the long term. How long will your Web site be sticking around? Ideally, like with any business, you want to build it to last without letting it fall behind and look dated. Relevancy to the current market is a big part of this, and if you are behind the times, you are probably behind your competitors. The technology that you use to build your Web site is inevitably going to change as the Internet advances, but your approach to relevancy should remain the same, incorporating new technologies as they arise. This is also a process you should develop over time. In the early days of the Web, frames were used to build sites, but that looks very outdated now. A few years ago, *splash pages* (introductory pages, mostly built in Flash, that provided no content or value to the user) were very popular. Today, they are discouraged. The Internet is an ever-changing entity, and if you're not persistent about keeping up with the times, you might fall by the wayside.

## ***Understanding the Search Engines: They're a Community***

Although dozens of search engines dot the Internet landscape, you'll be happy to hear there are really only a few you'll need to consider in your SEO planning. Each search engine appears to be a unique company with its own unique service. When people choose to run a search using Google, Yahoo!, Microsoft Live Search, Ask.com, or any of the others, they might think they've made a choice between competing services and expect to get varying results. But they'd be surprised to find out that under the surface, these seeming competitors are actually working together — at least on the data level.

Google's stated purpose is to "organize the world's information." When you think about the trillions of Web pages and multiple trillion words that exist, multiplying and morphing every day, it's hard to imagine a more ambitious undertaking. It makes sense, then, that not every search engine attempts such a daunting task themselves. Instead, the different search engines share the wealth when it comes to indexed data, much like a community.

You can see at a glance how this community works. Figure 1-3 shows how the major players in the search-engine field interact.



**Figure 1-3:** The Search Engine Relationship Chart depicts the connections between search engines.

**LEGEND**

- Supplies Receives Primary Search Results
- Supplies Receives Paid Results

*Chart courtesy of Bruce Clay, Inc.*

This search engine relationship chart includes all the major players. The arrows depict search results data flowing from supplying sites to receiving sites. Only four players, whose shapes are outlined — Google, Yahoo!, Microsoft Live Search, and Ask.com — are suppliers. They actually gather and provide search results data themselves. All of the non-outlined search



engines on the chart, including AltaVista, AOL, and the like, receive their search results data from some other source. The chart makes it clear that when you do a search on Netscape, for instance, the order of the results is determined by Netscape, but the indexed results are supplied by Google.

Bruce Clay's Search Engine Relationship Chart is also available online in an interactive Flash applet at

<http://www.bruceclay.com/serc/>

As the arrows depict, most of the search engines receive their data from one of these four sources. To further reduce the field, you can tell from the number of arrows coming from Google and Yahoo! that they feed the vast majority of other search sites. So in the world of SEO, you can feel pretty comfortable that if you're indexed in just two sites, Google and Yahoo!, you have a chance at ranking in most other search engines.

### ***Looking at search results: Apples and oranges***

One more thing to know about search results — there are two types. Figure 1-4 points out that a search engine can show these two different types of results simultaneously:

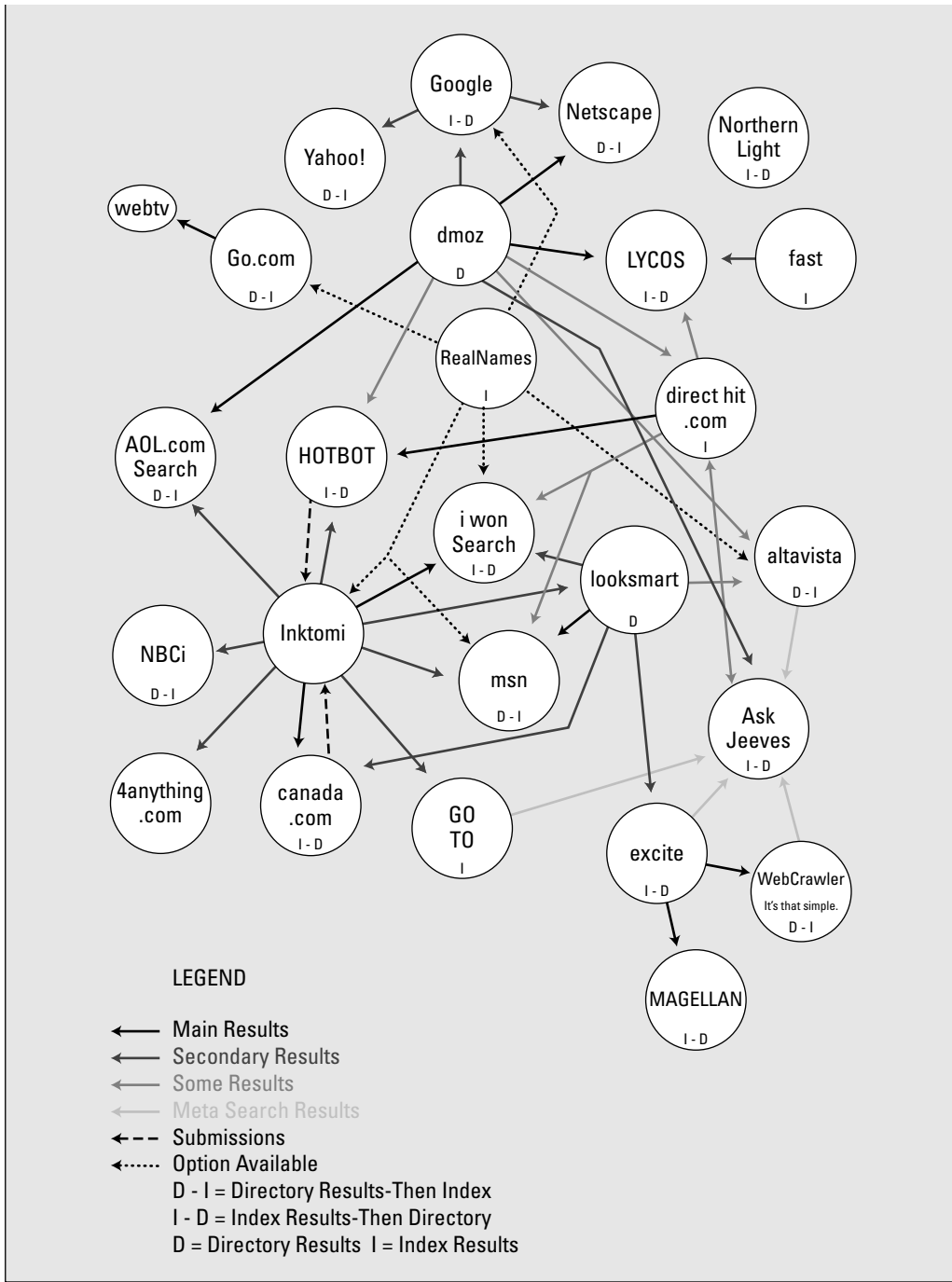
- ◆ Organic search results are the Web page listings that most closely match the user's search query based on relevance. Also called "natural" search results, ranking high in the organic results is what SEO is all about.
- ◆ Paid results are basically advertisements — the Web site owners have paid to have their Web pages display for certain keywords, so these listings show up when someone runs a search query containing those keywords. (For more on the whys and hows of paid results in greater detail, you can read about pay per click, in Chapter 5.)

## **A look back: Search engines a decade ago**

Bruce Clay first published his Search Engine Relationship Chart in 2000. Back then, there were more major players in the search game and things were, to say the least, somewhat cluttered. The chart had 26 companies on it: everyone from Yahoo! to Magellan to that upstart Google. Fifteen of those companies took their primary results from their own indexes; five of those supplied secondary results to other engines. Without a roadmap, it was an

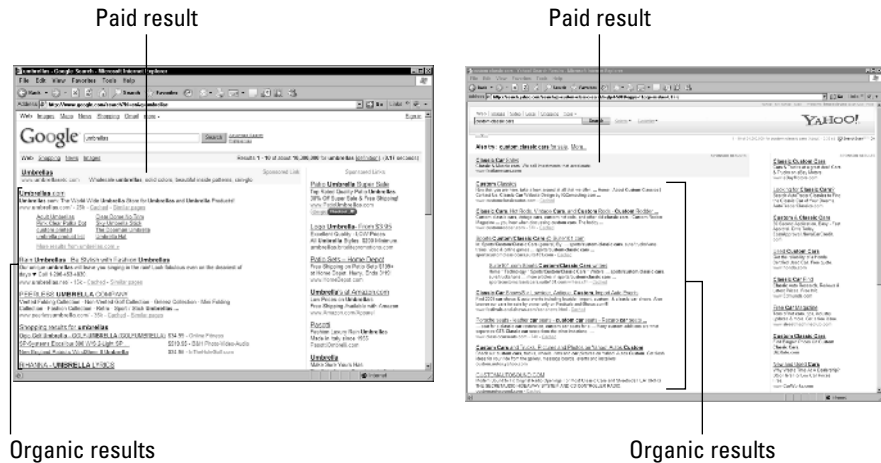
impossible task to keep it all straight. But over the years, things changed. What was once a cluttered mess is now a tidy interplay of a select group of companies. Here is an example of what the very first search engine relationship chart looked like:

Note: To view an interactive version of this chart online, check out [www.bruceclay.com/serc\\_histogram/histogram.htm](http://www.bruceclay.com/serc_histogram/histogram.htm).



On a search results page, you can tell paid results from primary ones because search engines set apart the paid listings, putting them above or to the right of the primary results, or giving them a shaded background, border lines, or other visual clues. Figure 1-4 shows the difference between paid listings and organic results.

**Figure 1-4:**  
A results page from Google and Yahoo! with organic and paid results highlighted.



The typical Web user might not realize they're looking at apples and oranges when they get their search results. Knowing the difference enables a searcher to make a better informed decision about the relevancy of a result. Additionally, because the paid results are advertising, they may actually be more useful to a shopping searcher than a researcher (remembering that search engines favor research results).

### *How do they get all of that data?*

Okay, so how do they do it? How do Google, Yahoo!, Ask.com, and Microsoft Live Search keep track of everything and pop up results so fast? Behold the wonder of technology!

Gathering the data is the first step. An automated process (known as *spidering*) constantly crawls the Internet, gathering Web-page data into servers. Google calls their spider the *Googlebot*; you could refer to them as *spiders*, *robots*, *bots*, or *crawlers*, but they're all the same thing. Whatever you call the process, it pulls in masses of raw data and does so continuously. This is why changes to your Web site might be seen within a day, or may take up to a few weeks to be reflected in search engine results.

In the second step, search engines have to index the data to make it usable. *Indexing* is the process of taking the raw data and categorizing it, removing duplicate information, and generally organizing it all into an accessible structure (think filing cabinet versus paper pile).

For each query performed by a user, the search engines apply an *algorithm* — basically a math equation (formula) that weighs various criteria and generates a result — to decide which listings to display and in what order. The algorithms might be fairly simple or multi-layered and complex.

At industry conferences, Google representatives have said that their algorithm analyzes more than 200 variables to determine search ranking to a given query. You're probably thinking, "What are their variables?" Google won't say exactly, and that's what makes SEO a challenge. But we can make educated guesses. (Same for Yahoo! and the others.)

So can you design a Web site that gets the attention of *all* the search engines, no matter which algorithm they use? The answer is yes, to an extent, but it's a bit of an art. This is the nuts and bolts of SEO, and what we attempt to explain in this book.

