

UNDERSTANDING HOW SEARCH ENGINES WORK

Search Engine Results

When a Web search is run with a major search engine (such as Google, MSN, Netscape, HotBot, alltheweb, etc.*), two types of results will typically be generated: sponsored sites and matching sites.

Sponsored Sites

The site or sites that appear first on your screen are usually sponsored sites. Sponsored sites are at the top because someone paid a fee for them to appear in response to search requests that contain specific words or phrases.

Typically, sponsored sites are operated by companies or individuals who have products or services to sell relating to the topic that is being searched. Most will have domain names that end with the suffix “.com.” Search engine companies generate income from displaying these sponsored sites. In most cases, sponsored sites are clearly labeled and separated from the “matching sites” generated by the engine.

Some search engines also have a section of sponsored sites that appear at the bottom of the Web page, and some also have them down the side of the page (such as this Google* search that has sponsored sites down the right side of the screen).

The screenshot shows a Google search for "brown recluse spiders". The search bar is at the top with the Google logo. Below the search bar, there are tabs for "Web", "Images", "Groups", "News", "Froogle", and "more". The search results are displayed below, with a "Web" tab selected. The results show "Results 1 - 10 of about 15,000 for brown recluse spiders (0.29 seconds)". The first result is a sponsored link for "Brown Recluse Information" from www.brown-recluse.com/. To the right of the main search results, there is a vertical column of "Sponsored Links" including "Brown Recluse Spiders" from www.brown-recluse-spiders.com, "Brown Recluse Spiders" from www.doyourownpestcontrol.com, "Brown Recluse Spiders" from www.Terminix.com, and "Camel Spiders" from www.captainbargains.com.

Matching Sites

Matching sites are sites that have been selected by a search engine as being relevant to a search request. They are listed in order of their perceived relevance to the search request. Their relevance may be determined automatically by the search engine, or a research staff may tweak the results to make sure that the most relevant sites appear at the top of the list. Web sites appearing in a list of “matching site” results have not paid fees to appear in the list. However, the site creators may have purposely manipulated their content to achieve a high listing from a search engine.

Matching sites may be operated by a variety of entities. For example, sites found by a search on a given topic might be run by: a federal, state, or local government agency; by a college or university, a manufacturer, a retailer, a school, church, club, or other organization; or by an individual person. The list of matching sites will often contain some or all of the sites that are listed as sponsored sites.

Search Engines

A variety of Web search engines are available for public use. They can vary significantly in the results that they generate for a Web search. These differences often happen because of the methods that they use to create their Web search listings.

Engines known as crawlers are designed to automatically seek out and analyze Web sites, utilizing special software to evaluate the content of those sites. Other than the software programming, there is essentially no human input involved in deciding how a crawler engine determines the relevance of a site to specific search terms.

Directory engines utilize a staff of researchers to analyze the content of Web sites and create relevant listings for common search criteria.

Hybrid engines utilize crawler search engine results in combination with direct human input to fine tune relevant listings for the more common search criteria. Google* is probably the most widely-used search engine today and is used for illustration in this publication. Google* is a crawler; it can be

accessed at <http://www.google.com>. Other search engines* (listed alphabetically) include:

AlltheWeb <http://www.alltheweb.com>

AltaVista <http://www.altavista.com>

AOL Search <http://search.aol.com>

Ask Jeeves <http://www.askjeeves.com>

HotBot <http://www.hotbot.com>

Lycos <http://www.lycos.com>

MSN Search <http://search.msn.com>

Netscape Search <http://search.netscape.com>

Teoma <http://www.teoma.com>

Yahoo <http://www.yahoo.com>

The home pages of some Internet portals (MSN, Netscape, Yahoo, etc.) have features for finding home telephone numbers and addresses of individuals and businesses. The operation of these features will vary with the provider. The White Pages features are generally very good at finding home telephone numbers and addresses. And once you have found the listing you are looking for, you can often print out a map and/or set of directions to get there. Finding business numbers with the Yellow Pages features seems to have a lower success rate but still works most of the time.

Yellow Pages and White Pages

**Search engines listed and used are for illustration purposes only and are not endorsed by the University of Kentucky, College of Agriculture or the Kentucky Cooperative Extension Service.*

Contact

Curt Judy, Cooperative Extension, Todd County, Rich Phelps, Agricultural Communications Services, and Darlene Taylor, Cooperative Extension, Logan County, January 2005.

Educational programs of Kentucky Cooperative Extension serve all people regardless of race, color, age, sex, religion, disability, or national origin.