1999

Search Tools: Skills and Strategies

Robin Schard
University of Miami School of Law, rschard@law.miami.edu

Search Tools: Skills and Strategies

Tired of wading through millions of Web sites when using Internet search tools? This guide helps you understand how search tools work and how to use them more efficiently.

By Robin C. Schard

Robin C. Schard, J.D., M.L.S., is Head of Public Services at the University of Miami Law School Library.

Historically, scientists used the Internet for point-to-point communication. As more data became available via the Internet, researchers needed an efficient way to locate relevant information. Early search tools, such as Gopher, Archie, Wais, and others, were not easy for the non-computer expert to use, requiring the searcher to remember different commands for each tool. With the development of the World Wide Web and hyperlinks, search tools became easier to use. Researchers could just point-and-click their way to interesting Web sites. This meant that even those without much computer experience could effectively use the Internet as a research tool.

Directories

Directories were one of the earliest Web search tools. They are basically catalogs of Internet sites, usually divided by subject. For example, Yahoo, one of the most popular directories, divides sites into broad categories and then subdivides the pages into narrower categories (Govern-

ment; Law; Elder Law). Another directory, Magellan, further enhanced these tools by using experts to evaluate and rank sites. Directories are still efficient search tools, but they do have some drawbacks. One disadvantage is in navigating the subject categories. Some Web pages do not exactly fit within existing categories, and directories must add more and more levels of subcategories to accommodate the ever increasing number of Web sites. The primary drawback is that they are limited in size, and the links can become outdated. Other search tools may be more comprehensive, but the restricted size of directories may also be a benefit when searching for basic information.

Search Engines

In order to address some of the drawbacks of directories, search engines were next developed. According to *The Dictionary of Multimedia Terms,* a search engine is "a program that locates and presents data, typically associated with a sizeable database."[1] Search engines actually go out and search the Internet periodically. It is all

done by computer. Search engines can be quite comprehensive and are often updated, depending on the program. Some search engines index the full text of a Web site, others only the title or Uniform Resource Locator (URL); some search engines digest the Web constantly, others less often. Another benefit of search engines is that the researcher no longer needs to rely on assigned categories but can pick the terms he or she thinks are important. Some search engines still include directories, allowing the researcher to select whichever tool works best.

Because of the size of the Internet and because search engines can index the full text of Web pages, the number of search results can be enormous. Depending on how the search is constructed, a results list in the millions is possible. Search engines have tried to address this problem through relevancy ranking. This allows a Web site to be ranked by how closely it matches the researcher's request. Typically, the ranking is determined by how often or how many search terms appear in the document (Web page). This means that documents containing (numerically) more search terms will appear near the top of the results list. Relevancy ranking does not end the problem of too many results, but it does lessen it. More sophisticated ranking programs are always being designed. Nevertheless, one of the main complaints about search engines remains that the searcher may have to wade through a list of hundreds of Web sites to find relevant information.

Other Search Tools

There are a couple of other search tools to consider when trying to locate information on the Internet. One is the metasearch engine. These tools search several search engines at once. One example is MetaCrawler (www.go2net.com/search.html), which searches Lycos, Infoseek, WebCrawler, Excite, AltaVista, Thunderstone, The Mining Co., Looksmart, and Yahoo. Metasearch engines should be used sparingly. If the main complaint about search engines is that the results list is too large, searching many engines simultaneously only compounds the problem. Metasearch engines can be used when looking for obscure information or unique terms.

Another tool, the minisearch engine, goes to the other extreme. Rather than search many sites, a minisearch engine only searches within one Web site. Despite their size, minisearch engines can be very valuable in searching complex or large Web sites. They do, however, tend to be very rudimentary, only searching terms or simple phrases.

Using Search Tools

My best advice concerning search tools is not to use them unless absolutely necessary. Even with all of the advancements in search technology, these resources can be unwieldy and unreliable. Executing a successful search still seems like a game of chance. If there is a quick, alternate method of locating the Web address for a site that will have relevant information, use that first, either by guessing the address, telephoning someone, or checking a paper source. If this is not possible, the following suggestions will provide some assistance.

The biggest mistake people make when using a search tool is in crafting their search. This breaks down into two errors. First, researchers do not carefully select search terms. A successful search will depend on the quality of the search terms. Creating on-line searches is not new to legal researchers, who have used commerical databases such as Lexis and Westlaw for years. Before beginning a search, the researcher should stop and think about the terms and phrases that will locate pertinent information while excluding irrelvant results. Although the researcher is not being charged for each transaction on the Internet (assuming we are not discussing Lexis and Westlaw or some other commercial database on the Internet), it is still vital to plan the search beforehand. This is particularly important on the Internet, because the search tools are not yet terribly sophisticated and therefore it will prevent wandering aimlessly on the Web. So, when planning to search the Internet, as with all things, remember, "prior proper planning prevents poor performance."

The second common mistake is not using the advanced search features most tools now provide. Most search tools have two search designs: one for the novice searcher and one for the advanced searcher. Sophisticated searchers are those who understand how to construct searches with connectors or Boolean logic, or who have been trained to search on-line databases. The advanced features give the researcher better control over the search. For example, rather than browse through 7,660 sites from HotBot when looking for "Medicare manuals" (the first government site is number 10, and it is in Spanish), it would be easier to look at the 450

advanced search results by limiting the search to English-language sites including the words "Medicare" and "manual" and "HCFA" with a government domain name (.gov). (The HCFA site is the first result, and the Medicare Program Manuals is the second.) It does, however, take a little more time to use these features. While there is some standardization, each tool has its own toolbox of advanced search features. These features are usually detailed in the "help" or "advanced" section of the resource. I would not recommend memorizing the features for every search tool, but learning one or two sources very well will save time in the end. Table 1 details some, but certainly not all, of the advanced features for a few search tools.

In addition to constructing a satisfactory search, the researcher must also decide which search tool to use as a starting point. Generally speaking, it is best to begin with the smallest tool and expand as necessary. For instance, if you are searching for the Wyoming statutes, it is much faster to go to the State Cases and Codes section of Findlaw, which is a directory specifically limited to legal material, than to search Infoseek for the same material. (An Infoseek search for Wyoming statutes has a results list of 1,536,118, and the direct link to the current statutes does not appear in the first 20 hits.) Even when searching for general infor-

mation, the researcher may wish to start with less comprehensive resources. Directories may, indeed, be better starting points than search engines. If, for example, the searcher needs a Canadian postal code, Yahoo's Reference section has postal information, but an Excite search does not reveal the direct link to the Canadian Post in the first 20 results. As with understanding advanced search features, learning a few search tools really well will greatly aid the researcher in selecting a starting point. A person using search tools should know at least one directory (e.g., Yahoo), one comprehensive search engine (e.g., HotBot), and a subject-specific search tool, if applicable (e.g., Lawrunner).

Of course, educating oneself in the use of search tools is not a one-time experience. Search technology is changing just as rapidly as other technology. As new tools and new features are introduced, Internet researchers must update their knowledge. The sophisticated searcher will need to read new literature about the Internet and notice changes in the design or layout of favorite sources. Some suggested current awareness resources are:

- Internet World Online (www.iw.com): Evaluates all types of Internet innovation, including search capabilities
- PC Magazine Online (www.zdnet.com/pcmag): On-line publication that often describes updates to search

tools, including ranking many of the possible resources. (The December article looked at 82 different sites: http://www.zdnet.com/pcmag/features/websearch98/intro.html)
- Search Engine Watch (www.searchenginewatch.com): Web site that evaluates major search engines. It also provides tips on how best to use these resources and links to other sources that evaluate search tools. Great site for any possible question concerning search tools

Conclusion

Overall, the main consideration in using search tools is to become familiar with these resources before you have to use them. Once the need for information arises, it is important to treat the search just like any other electronic search. Prepare before going on-line; decide which source is most likely to have the information, and do not waste too much time before asking an expert for assistance. If you use these tips, your Internet searches should be more rewarding and less stressful.

Endnote

1. Brad Hansen, THE DICTIONARY OF MULTIMEDIA: TERMS AND ACRONYMS, 194 (1997).

Table 1. Selected Search Tools and Advanced Features

Search Tools	Connectors*	Proximity*	Truncation	Other
Findlaw (www.findlaw.com) Directory of legal resources that may be browsed by general subject or searched by word; an excellent place to find basic, primary legal material such as codes, court decisions, or directories.	AND OR NOT	NEAR (ranked by proximity of terms)	* = wildcard **=variants of a word. E.g., fly** finds flying, flown, flew	Allows user to search for specific legal subjects, such as cases, codes, tax, etc.
Lawrunner (www.lawrunner.com) Takes advantage of Altavista's powerful search engine but is preset to look for legal sites. Searches can also be restricted to several selected legal and governmental sites.	AND (&) OR (l) AND NOT (!) (+) include (-) exclude	NEAR (~) (terms within 10 words of each other) " " for exact phrases	* = wildcard	text: title: url: domain: Refine—like a thesaurus Case sensitive
Altavista (www.altavista.com) Very large and easy to use.	See Lawrunner	See Lawrunner	See Lawrunner	See Lawrunner
Excite (www.excite.com) Excite does have some general reference material, and are sites arranged by subject. Power search allows user to restrict search through dropdown boxes.	AND OR AND NOT (+) include (-) exclude	" " for exact phrases		More Like This - searches for similar pages Program searches for concepts rather than the specific terms
HotBot (www.hotbot.com) Powerful and easy to use. It supports some interesting search features, including searching by date, domain, or media type.	AND OR NOT (+) include (-) exclude	" " for exact phrases	* = wildcard Check box for word stemming	Allows user to restrict search through dropdown boxes Case-sensitive search when you insert capital letters
Infoseek (www.infoseek.com) Good search format.	(+) include (-) exclude	" " for exact phrases		Search within these results—refines search further Advanced search provides convenient forms Case-sensitive search when you insert capital letters

Table 1. Selected Search Tools and Advanced Features (*continued*)

Search Tools	Connectors*	Proximity*	Truncation	Other
Magellan (magellan.excite.com) Subject-oriented directory that allows searching within the directory.	AND OR NOT (+) include (-) exclude	" " for exact phrases		Find Similar—locates similar pages Many of the sites are evaluated and ranked
Yahoo (www.yahoo.com) Subject-oriented directory that allows searching within the directory. If your search is unsuccessful, it will default to an Inktomi search.	AND OR (+) include (-) exclude	" " for exact phrases	* = wildcard	[t:] restricts searches to title [u:] restricts searches to URL Pull-down menu for date and other restrictions

* Boolean connectors and proximity terms, such as "AND" and "NEAR," should all be in capital letters.