



LLRX[®]
com

More than 300 legal software reviews

Front Page Bookstore Archives About LLRX Subscribe Comments

SEARCH **GO!**

Navigation

Features

ResearchWire

Update to *Search Engines Compared*

By Diana Botluk

[Diana Botluk](#) is a reference librarian at the Judge Kathryn J. DuFour Law Library at the Catholic University of America in Washington, D.C., and is the author of the [The Legal List: Research on the Internet](#). She teaches legal research at [CAPCON](#), [Catholic University Law School](#), and the [University of Maryland](#). Take a class with Diana! [Here's how...](#)

Published July 15, 2000



Click Here for a [PDF](#) format of this article

General Web search engines are getting smarter. Yes, I know, only living creatures get smarter. Computers don't think, they only do math. But in the case of search engines, this year they're doing it better. Many of the general Web search engines have added a new feature that attempts to anticipate what it is the researcher wants, sort of the way librarians sometimes have to guess at what it is their patrons really want, instead of what they ask for. Have you seen the Web site Ask Jeeves (<http://www.askjeeves.com>). This exemplifies the concept. Researcher types in search request and AskJeeves replies with many possible questions which might actually represent what researcher is looking for. For example, I typed in campbell soup and asked Jeeves. Before Jeeves listed my Web search results he came back and told me he found answers to the following questions, and listed several questions like "what is the company Web site?" or "where can I find stock quotes?" Clicking on a questions retrieves its answer.

So what does that have to do with general Web search engines? It seems this is the hot new feature for the year 2000, and more of the general Web search engines are incorporating anticipated questions into their Web results. Of the six search engines I included on the features comparison, AltaVista and Excite suggested questions I might want answered. While Go.com did not suggest questions to be answered, it did suggest possible new searches I might want to try. And both Lycos and HotBot told me what other searches people did who did the same search as me, sort of like Amazon's "people who read this book also read...."

Although general Web search engines are getting easier to use, sometimes locating relevant information is still like looking for a needle in a haystack. General search engines are not always the proper tool to use when trying to locate specific information. It is often better to locate a database likely to contain the desired information and search for it within that database.^{*} However, sometimes that isn't possible. And sometimes we want to be more general, casting a wider net to see what we come up with. At these times, we may choose to use a general search engine. There are so many general Web search engines that often we don't know which one to choose. The chart that follows will compare the features of six popular search engines. In this way, researchers can see at a glance which search engine offers which features, and be more informed as to their choices. Ultimately, a researcher armed with this knowledge will be able to pick the best tool for the particular job.

AltaVista

AltaVista provides a lot of search construction options for sophisticated searchers, and therefore has long enjoyed favor by information professionals. Besides traditional Boolean search options, AltaVista has many field restrictors, and also has an interesting forced phrase searching feature. In a basic AltaVista search, the search engine treats two terms together that are statistically most likely to be a phrase as a phrase, even without quotes. This changes in AltaVista's advanced search, where it depends on which search box you use. In the Boolean search box, all terms are treated as a phrase unless there is a connector between them. In the Sort by search box, terms are treated individually, and a phrase must be designated by quotes. Note that AltaVista not only supports truncation, but also internal wildcards when and asterisk (*) is used.

AltaVista translates its results into several different languages, which can come in handy when you run across a page in a language you don't understand. The translation is not precise, but it can be good enough to get the gist of what is on the page. AltaVista also has a special photo and media finder. The photo finder returns a results list complete with photo thumbnails. Like Google, AltaVista uses RealName Internet keywords to link to the home pages of companies or brand names. Additionally, on its results display AltaVista has a feature which prompts questions that might be in the minds of researchers, and provides links with the answers to these questions.

Excite

Excite provides a few more options than Google for creating a detailed search, but the beauty of Excite is its concept searching. Excite is a good place to cast a wide net over the Internet ocean. If you type a word in the search box, not only do you search for that word itself, but also forms of the word, synonyms of the word, and other words that are highly statistically related to that word. Thus, Excite is a good search engine to give yourself a little leeway, when you don't necessarily want the search terms to be as precise as they can be. However, if you use a Boolean operator in the search string, remember that it turns the search into an exact keyword search, eliminating any concept

searching for word variations, synonyms, etc.

Go

When Infoseek and Disney joined forces, the portal Go.com was created. Thus the search engine at Go.com is the old Infoseek search engine, and still employs many of the same features. Go.com supports Boolean searching and allows searching with many restrictors. Additionally, the advanced search provides a variety of specialized search engines for finding specific information. A popular option here is the Companies search, which results in company capsules and home page urls, with options for stock and press information. A Go.com search results page provides many options besides a list of relevant search results. It offers possible related searches as well as the option of allowing the search engine to create a new search automatically based on a specific page. It also provides the opportunity to search in levels, by performing one broad search first, then narrowing it down by searching again within the given results, rather than searching the Web all over again. Its translation option allows the translation of any page.

Google

This is the simplest search engine to use, but in some ways it is dangerously simple. Google provides very few options for searchers to construct a detailed search. In fact, just about the only options are a minus sign to exclude terms and parentheses to force terms to be searched as a phrase. BEWARE! Google has only inclusive searching with no option for alternative searching besides starting another search. What does that mean?

Let's say I wanted to search for pages that say either attorney or lawyer. With most search engines you would have an option of searching for either term at the same time. They would search attorney OR lawyer automatically, making alternative searching the default. Or you may have to physically insert the Boolean connector OR between the words to force the alternative search.

With Google, typing attorney lawyer in the search box will only search for pages that say both attorney and lawyer. Google does not support the Boolean connector OR. The only way to search for pages containing either attorney or lawyer is to perform two searches.

Google's results list provides a similar pages feature that allows the computer to construct automatically a new search for pages similar to a specified page - a good option if you find one good page and want more like it. It provides an option for highlighting search terms on the resulting page, something not often seen in Web search engines. Finally, Google makes use of RealNames or Internet keywords, which means results on the list with that mark will link directly to the company or brand name home page. This can be extremely useful. Suppose you want to search for The Walt Disney Company home page. Simply typing disney in any search engine's search box is likely to result in hundreds of thousands of hits, listing every page where the word Disney appears. This is not an efficient way to search for the company's home page, unless the search engine uses the RealName mark to distinguish that hit as the one that will deliver the home page. On Google, these results are listed first.

HotBot

HotBot's form interface still makes it a favorite among many researchers. It allows you to build a fairly sophisticated search without having to

remember connectors or restrictors, since they all appear right there on the form and you can choose and click on what you want. The best way to use HotBot is to skip the first page and go directly to the Advanced Search page, which provides many more search options.

HotBot results employ direct hit technology, which provides a list of the top ten most popular links for any given search. The theory behind direct hit is that the sites that people go to most based on a given search are also likely to be the most relevant sites for that search. Another interesting item on the search results page is a list of possible related searches from which to choose to have the search engine automatically perform another search.

Lycos

Lycos also supports Boolean searching, and by far has the most extensive options for proximity searching as any search engine on the Web. Lycos, like Go.com, provides the option of searching by levels, where you can search within a previous set of results. It will also offer suggested searches following your initial search. Lycos' advanced search provides a variety of specialized search engines for locating specific information. An automated tracking feature at Lycos allows users to register and have their searches updated automatically. Lycos' results page offers a popular links region where the most popular links for certain searches will be distinguished from regular results.

****For permission to distribute this search engine comparison chart for educational purposes, please contact [Diana Botluk](#)****

Comparison of Major Web Search Engines

		AltaVista www.altavista.com	Excite www.excite.com	Go go.com	Google www.google.com	HotBot www.hotbot.com	Lycos www.lycos.com
	default searching	alternative	alternative	alternative	inclusive	inclusive	inclusive
	inclusion of words	simple search: plus sign (+) (simple search); AND or plus sign (+) (advanced search)	AND or plus sign (+) "must have" on the advanced search form	plus sign (+) or use must on the advanced form	automatic	choose "all the words" on the search form, the "must contain" option on the search form, or use AND or the plus sign (+) with the Boolean expression option	AND, plus sign (+) or choose "all the words" on the advanced search form

<p>alternative words</p>	<p>OR in an advanced search</p>	<p>OR</p>	<p>all words are searched alternatively unless otherwise specified</p>	<p><i>not supported</i></p>	<p>choose "any of the words" on the search form or use OR with the Boolean expression option</p>	<p>OR or choose "any of the words" on the advanced search form</p>
<p>proximity</p>	<p>NEAR finds terms occurring within 10 words of each other in an advanced search</p>					<p>NEAR: terms within 25 words of each other</p> <p>NEAR/#: terms within specified number of words of each other</p> <p>ONEAR: same as NEAR, but with words in specified order</p> <p>ONEAR/#: same as ONEAR, but with words in specified order</p> <p>FAR: terms appear at least 25 words apart</p> <p>FAR/#: terms appear at least specified number of words apart</p> <p>OFAR: same</p>

Search Language							as FAR, but with words in specified order OFAR/#: same as OFAR, but with words in specified order BEFORE: finds pages where the first term appears before the second
	exclusion of words	minus sign (-) (simple search); NOT or minus sign (-) (advanced search)	NOT or minus sign (-)	minus sign (-) or should not on the advanced form	minus sign (-)	Choose "must not contain" on the search form; NOT or the minus sign (-) with the Boolean expression option	NOT or minus sign (-)
		quotation marks ("") or hyphens (-) between words <i>FORCED PHRASE SEARCHING:</i> <i>basic search will assume 2 terms together that are statistically most</i>	quotation marks ("") In Advanced Search, multiple words in a single keyword box are automatically treated as a phrase.	quotation marks ("") or phrase on the advanced form proper names with initial capital letters are searched as a phrase even without	quotation marks ("")	quotation marks (""), or choose "exact phrase" on the search form	quotation marks ("") ADJ: words must appear next to each other ADJ/#: terms must appear exactly

<p>phrases</p>	<p><i>likely a phrase will be searched as a phrase, even without quotes</i></p> <p><i>advanced search: terms in the Boolean box are treated as a phrase unless there is a Boolean connector; terms in the sort by box are treated separately unless enclosed in quotes</i></p>		<p>quotes; commas between capitalized names force them to be searches separately</p>			<p>specified number of words apart</p> <p>OADJ: same as ADJ, but words must be in specified order</p> <p>OADJ/#: same as ADJ/#: but words must be in specified order</p>
<p>case sensitivity</p>	<p>lower case search searches are case insensitive; upper case searches force case sensitivity</p>	<p>not case sensitive</p>	<p>lower case searches are case insensitive; capitals forces case sensitivity</p>	<p>not case sensitive</p>	<p>lower case searches are case insensitive; mixed upper/lower case searches are case sensitive</p>	
<p>truncation/wildcard</p>	<p>asterisk (*) for truncation or internal wildcard – must appear after at least 3 characters and can replace 0 to 5 characters</p>	<p>not necessary with concept searching</p>			<p>truncate words with an asterisk (*) or choose "enable word stemming" on the form</p> <p>asterisk (*) is a wildcard that can replace any number of characters; question mark (?) is a wildcard that can replace only</p>	

						one character; either can be used anywhere in a word.	
	nesting	parentheses ()	parentheses ()			parentheses () with the Boolean expression option	
	date	form based date restriction available with advanced search				restrict by date on the search form	
	language	yes	yes		yes		yes
	location		yes (advanced search)	yes (advanced search)		restrict by location on the search form	
	media type	Special media finder. Image finder returns thumbnails in results.				restrict by media type on the search form; or by using feature:	
	title (searches specific word or phrase in a page's title field)	title:		title:		title:	use the form option
	url (searches specific word or phrase in a page's url)	url:		url:			use the form option
		anchor: specified word in the text of a hyperlink	domain type (advanced search)	link: site:	link:	domain: pages within the specified	

**Search
Restrictors**

<p style="text-align: center;">Other Search Features</p>	<p>Special photo and media finder</p> <p>related searches can be performed by clicking on related pages on search results page</p> <p>basic search ranks by relevancy; advanced search users MUST use the sort by box to control ranking preferences</p> <p>punctuation other than mentioned in search language and restrictors is read as a word separator, or blank space</p>	<p>using concept searching, now called <i>intelligent search</i>. Concept searching will search for other word forms, synonyms, and related terms in addition to the keyword typed into the search statement</p> <p>use of Boolean operators (AND, OR, NOT) forces keyword (exact term) searching rather than concept searching</p>	<p>set of results by using the pipe symbol (), or clicking the search within results box</p> <p>allows a new search to find pages similar to a chosen result</p> <p>offers possible related searches by clicking on a generated list of similar searches</p>	<p>searched</p>	<p>makes advanced searching easy and eliminates the need to investigate appropriate search language</p> <p>direct hit technology displays top ten most popular pages from given results at the top of the results display (not available with all searches)</p> <p>a list of possible related searches from which to choose appears on the results page</p> <p>allows a second level of searching within a given set of results</p> <p>searching for the person on the search form searches</p>	<p>previous set of results</p> <p>offers suggested searches after an initial search</p>
---	--	---	--	-----------------	--	---

					automatically for address and e-mail information	
Results Display	<p>display reveals short summary, url, file size, page date and language</p> <p>word count reveals the number of times each search term appears</p> <p>translation option allows translation of any page</p> <p>site compression in basic search shows only one page per web site; site compression is not automatic in advanced search, but can be turned on by a check box</p> <p>image finder results are thumbnail images</p> <p>results are limited to 200, but advanced search users can change the url to go</p>	<p>display reveals url, relevancy score, and summary</p> <p>results show hits in directory first, then web, then news</p> <p>user can sort results by site or relevancy</p> <p>similar search feature allows a related search</p> <p>Quick Results" provides fast answers to the most popular questions.</p>	<p>display reveals relevancy score, url, date, summary and file size</p> <p>first shows matching directory topics, then web pages</p> <p>translation option allows translation of any page</p> <p>users can choose to hide summaries</p> <p>users can sort by date</p> <p>results clustering prevents all top hits from being from same site; results clustering can</p>	<p>display reveals url, summary and file size</p> <p>highlights search terms on results pages</p> <p>similar pages feature allows a related search</p> <p>users can set number of results per page</p> <p>RealNames (RN) mark designates the result will link directly to company or brand name home page</p>	<p>full descriptions include summary, url, file size and page date</p> <p>search form allows users to set results to full descriptions, brief descriptions, or urls only</p> <p>results clustering prevents all top hits from being from same site</p> <p>search results also offer results from the Hotbot directory</p> <p>Hotbot displays the top ten most popular sites first; if</p>	<p>results include summary & url</p> <p>results are organized into web sites, news articles, shopping, and the most popular links for given search terms</p> <p>results also match directory categories</p> <p>suggests refine terms</p>

	<p>beyond 200</p> <p>ASK ALTA VISTA: in basic search, when a common English sentence, phrase or question is entered, AV prompts that it knows the answer to the question. users choose from a lits of questions and AV finds the answer in a special database</p> <p>AV may also prompt a better way of entering a query if many others have used that query before</p> <p>a list of RealName (RN) internet keywords can be clicked to</p>		<p>be turned off with ungroup results</p>		<p>you change the default search settings, the top ten results may appear as a link rather than being immediately listed</p> <p>a refine option on the search page suggests new or alternative terms to add to the search</p>	
Subject Directory	yes	yes	yes	yes	yes	yes
Other Special Features	<p>when result is from a web site of an organization in AV's company/organization database, AV allows users to jump to a company fact sheet</p>	<p>provides a variety of specialized search engines for specific information</p>	<p>advanced search provides a variety of specialized search engines for specific information</p>		<p>results page offers a variety of options to locate information in alternative ways</p>	<p>automatic tracking feature lets users register and have their searches updated automatically</p> <p>advanced search provides a variety of</p>

						specialized search engines for specific information
--	--	--	--	--	--	---

Footnote

* See [Strategies for Online Legal Research: Determining the Best Way to Get What You Need](#), LLRX, April 3, 2000; [Exposing the Invisible Web](#), LLRX, October 1, 1999.



[[LLRX Front Page](#) | [Bookstore](#) | [Archives](#) | [About LLRX](#) | [Subscribe](#) | [Comments](#)]

Copyright © 1996-2000  Law Library Resource Xchange, LLC. All rights reserved.