


# Cheat Sheet #1

## General Search Engine Features Comparison Chart (Limiters in Cheat Sheet #2)

Features shared by all SIX SEARCH ENGINES	Examples:
• Default <b>AND</b> between words	No need to type <b>AND</b>
• Double quotes " " makes a phrase search, exactly as typed	<b>"search engines"</b>
• <b>OR</b> (must be capitalized) to allow any or either word or phrase in quotes. Yahoo, MSN, Gigablast, and Exalead offer full Boolean with AND NOT, ( )	<b>web OR internet</b> <b>"search engines" OR "subject directories"</b>
• Common words ignored + before forces them to be searched " " words enclosed in a phrase will be searched	Little consistency what words are common. Look at search results to find out what was searched. <b>+all +in +a day's work</b> <b>"all in a day's work"</b>
• <b>Minus (-)</b> excludes. Cannot be used in combination with full Boolean searches with AND NOT, OR, ( )	<b>pink -floyd</b> <b>"search engine" tutorial -site:com</b>

Search Differences	Google	Yahoo! Search	Ask.com	MSN Search	Gigablast	Exalead
Size, indexing limit	<b>Huge</b> Probably the largest • Indexes only first 101 KB of a web page and about 120 KB of PDFs	<b>Huge</b> 22 billion "web objects" • Indexes first 500 KB of a web page	<b>Moderate</b> 2+ billion	<b>Huge</b> 5+ billion	<b>Moderate</b> 2+ billion	<b>Small, beta</b> New, growing, 4 billion
Ranking	<b>Popularity</b> Main factors: • links to a page • "importance" of linking pages • Subjective PageRank™ value of 1-10 • Word order, proximity	<b>Relevancy</b> In Yahoo Mindset (beta), slider to vary: • shopping, commercial • research, academic, informational, non-commercial mindset.research.yahoo.com	<b>Same-subject popularity</b> "based on the number of same-subject pages that link to it"	<b>Relevancy</b> In Search Builder, sliders to vary: • recently updated ↔ static • very ↔ less popular • approx. match ↔ exact match	<b>Relevancy</b>	<b>Relevancy</b> Date sorts in Advanced Search (oldest, newest)
Clustering results from same site	• 2 results from a site • Link to show more	• 1 results from a site • Link to show more	• 2 results from a site • Link to show more	• 2 results from a site • Change or turn off in Settings	• 1 result from a site • Turn off clustering in Adv Srch	• 1 result from a site • More by clicking folder
Clustering by content		• Related searches suggested when other Yahoo searches have searched your topic	<b>Teoma:</b> • "Refine" – ways to narrow search, adds terms • "Resources" – link pages <b>Ask Jeeves:</b> • Narrower search terms • Broader search terms		• "Giga Bits" show % of, and retrieve, pages with recurring terms • Reference pages – link pages • Related searches	• Related searches box
For-fee sites included	Identified sponsor sites, not in results	Identified sponsor sites Mixed in results: • Pay-for-inclusion • Pay-for-position	<b>Ask Jeeves:</b> Identified sponsor sites Mixed in results (few): • Pay-for-inclusion • Pay-for-position	Identified sponsor sites Mixed in results: • Pay-for-inclusion • Pay-for-position	None	None

Search Differences	Google	Yahoo! Search	Ask.com	MSN Search	Gigablast	Exalead
<b>Boolean beyond OR</b> Operators must always be capitalized. In most search engines that allow Boolean, you must enclose words joined by OR in ( ), and type AND when implied	OR, - <ul style="list-style-type: none"> <li>• <b>pets cat OR dog -collars</b></li> </ul>	OR, AND, NOT, AND NOT, ( ) Cannot use - with operators <ul style="list-style-type: none"> <li>• <b>pets AND (cat OR dog) AND NOT collars</b></li> <li>• <b>pets AND (cat OR dog) NOT collars</b></li> </ul>	OR, - <ul style="list-style-type: none"> <li>• <b>pets cat OR dog -collars</b></li> </ul>	OR, AND, NOT, AND NOT, ( ) Cannot use - with operators <ul style="list-style-type: none"> <li>• <b>pets AND (cat OR dog) AND NOT collars</b></li> <li>• <b>pets AND (cat OR dog) NOT collars</b></li> </ul>	OR, AND, AND NOT, ( ) <ul style="list-style-type: none"> <li>• <b>pets AND (cat OR dog) AND NOT collars</b></li> </ul>	OR, AND, NOT, AND NOT, ( ), NEAR Cannot use - with operators <ul style="list-style-type: none"> <li>• <b>pets AND (cat OR dog) AND NOT collars</b></li> <li>• <b>pets AND (cat OR dog) NOT collars</b></li> <li>• <b>pets NEAR collars</b></li> </ul>
<b>Other search control terms</b>	~ for synonyms <ul style="list-style-type: none"> <li>• <b>~FAQ finds FAQ, help, manual, etc.</b></li> </ul>			<b>prefer:</b> before a term specifies the term without requiring it <ul style="list-style-type: none"> <li>• <b>cow prefer:mad</b> finds cow, preferably mad</li> </ul>		<b>OPT</b> before a term specifies the term without requiring it <ul style="list-style-type: none"> <li>• <b>cow OPT mad</b> finds cow, preferably mad</li> </ul> <b>Phonetic search</b> in Adv. Srch. finds sound-alikes <ul style="list-style-type: none"> <li>• <b>movy staar</b> would find movie star</li> </ul>
<b>Stemming</b>	<ul style="list-style-type: none"> <li>• Stems some words</li> <li>• + turns stemming off</li> <li>• No stemming within phrases in quotes</li> </ul>	<ul style="list-style-type: none"> <li>• No</li> </ul>	<ul style="list-style-type: none"> <li>• No</li> </ul>	<ul style="list-style-type: none"> <li>• No</li> </ul>	<ul style="list-style-type: none"> <li>• No</li> </ul>	<ul style="list-style-type: none"> <li>• Stems some words if turned on in Preferences.</li> <li>• + turns stemming off</li> </ul>
<b>Truncation</b>	<ul style="list-style-type: none"> <li>• None.</li> </ul>	<ul style="list-style-type: none"> <li>• None</li> </ul>	<ul style="list-style-type: none"> <li>• None</li> </ul>	<ul style="list-style-type: none"> <li>• None</li> </ul>	<ul style="list-style-type: none"> <li>• None</li> </ul>	<ul style="list-style-type: none"> <li>• * allows any ending</li> </ul>
<b>Hyphen</b>	<ul style="list-style-type: none"> <li>• Searches as phrase, two words, one word</li> </ul>	<ul style="list-style-type: none"> <li>• Ignored</li> </ul>	<ul style="list-style-type: none"> <li>• Searches as hyphen and as two words with space</li> </ul>	<ul style="list-style-type: none"> <li>• Ignored</li> </ul>	<ul style="list-style-type: none"> <li>• Ignored</li> </ul>	<ul style="list-style-type: none"> <li>• Ignored</li> </ul>
<b>Cached pages</b>	<ul style="list-style-type: none"> <li>• Cached links</li> </ul>	<ul style="list-style-type: none"> <li>• Cached links</li> <li>• Link to Internet Archive Wayback Machine in cached page</li> </ul>	<ul style="list-style-type: none"> <li>• Cached links</li> <li>• Thumbnails of pages when </li> </ul>	<ul style="list-style-type: none"> <li>• Cached links</li> </ul>	<ul style="list-style-type: none"> <li>• Archived copy</li> <li>• Date updated</li> <li>• Older Copy = Link to Wayback Machine</li> </ul>	<ul style="list-style-type: none"> <li>• Thumbnail Images of pages as crawled</li> </ul>
<b>Search size limit</b>	<ul style="list-style-type: none"> <li>• 32 words</li> </ul>			<ul style="list-style-type: none"> <li>• 10 words</li> </ul>		
<b>Other Unique Search features</b>	<ul style="list-style-type: none"> <li>• * whole-word wild card: <b>"it's * * * day"</b> <b>it's * * * day</b></li> <li>• <b>define:[term or concept]</b> finds definitions on the web: <b>define:internet</b></li> </ul>	<ul style="list-style-type: none"> <li>• Yahoo! Directory Links</li> <li>• Creative Commons search in Adv. Srch (find sharable software)</li> </ul>	<ul style="list-style-type: none"> <li>• <b>what is Internet define Internet smallest moon</b></li> <li>• <b>[Web Answers]</b> finds answers in web pages, at top of results</li> <li>• <b>Quick Definitions</b> from dictionary at top of results</li> </ul>	<ul style="list-style-type: none"> <li>• Search Builder</li> </ul>	<ul style="list-style-type: none"> <li>• DMOZ Open Directory links in results</li> <li>• Government sites database (search tab)</li> </ul>	<ul style="list-style-type: none"> <li>• Complex results screen, many options, options vary with search results.</li> <li>• DMOZ Open Directory links</li> </ul>