

## Searching the Web: Toward Maximizing Relevance

Mark A. Spasser, PhD, MLS, *Jewish Hospital College of Nursing and Allied Health, St. Louis, Missouri*

Searching for information on the Web is paradoxically easy but frustrating. There is an increasingly vast amount of high-quality information available on the Internet, but finding it can seem like groping through the world's largest haystack for the proverbial needle. This review discusses some of the most basic and important methods and techniques, (table 1) applicable to almost any search using almost any search engine, whose thoughtful application ensures that the information found is both high-quality and relevant.

### REPRINT REQUESTS:

Mark A. Spasser, PhD, MLS  
Chief, Library and Information  
Services/Associate Professor  
Jewish Hospital College of Nursing  
and Allied Health  
306 S. Kingshighway Blvd.  
MS:90-30-625  
St. Louis, MO 63110-1091  
Telephone: 314-454-8171  
Email: mas1200@bjc.org

### KEYWORDS:

Internet; Search engines; Queries

### RECEIVED:

August 1, 2002

### REVISED AND ACCEPTED:

August 13, 2002

### Clinical Medicine & Research

Volume 1, Number 1: 69 - 70  
©2003 Clinical Medicine & Research  
www.mfldclin.edu/clinmedres

### TYPES OF SEARCH ENGINES

Centralized - Google, Teoma, and AlltheWeb

Specialized/expert - Pubmed (Medline), AllLaw, MedHunt

Meta search engines - Meta engines are becoming less important due to centralized search engines indexing the Web so comprehensively, and as a result will not be covered in this review

### GUIDING PRINCIPLES

*Be specific:* Centralized search engines, such as Google and AlltheWeb, work better when you use multiple keywords that are unlikely to pop up on irrelevant pages. Also, of course, use specialized search engines whenever possible.

*Evaluate retrieval:* Listing length of candidate documents returned by queries submitted to search engines is directly related to amount of evaluation time needed.

*Gauge time:* Whereas actual search time in getting candidate listings from search engines is relatively fast, document-by-document download and review is the time consuming part of the process.

*Read instructions:* Analogous to database interfaces, every search engine has its own syntactical peculiarities; entering a query properly can mean the difference between finding what you're after and being deluged by irrelevant links.

*Go Boolean:* Most sites permit the construction of queries using the AND, OR, and NOT Boolean operators (use parentheses if you use more than one Boolean term). Also the + and - implicit operators, are almost universally supported.

*Be exacting:* Surround words with quotation marks to search them as a phrase. The utility of this widely supported search feature to ferret out only the most relevant information can't be overstated.

*Get advanced:* Many engines come with an advanced search page offering additional query types and sorting options, which are especially helpful when a simple search bombards you with pages of links.

*Be steadfast:* If the first site is not successful, try another; bookmark several and try new contenders from time to time. Each engine has potentially complementary strengths and limitations.

*Save pages:* Once documents are located, keep track of them; at the very least bookmark them.

Table 1. Useful search features common to most search engines.

Search Feature	Example	Rule
Implied Boolean Operators		
+	+bipolar +depression	The + symbol instructs the search engine to return only pages that <b>include</b> that word (or words).
-	+bipolar +depression -psychotic	The - symbol instructs the search engine to return only pages that <b>exclude</b> that word (or words).
Power of Boolean Searches		
AND	bipolar AND depression	AND indicates that only those web pages having <b>all words</b> in them will be retrieved (narrows search).
OR	(bipolar OR manic) AND depression	OR indicates that web pages having <b>any term</b> will be retrieved (broadens search). OR is the operator used to construct searches using synonyms.
NOT/AND	bipolar AND depression AND NOT psychotic	NOT/AND indicates that web pages having the term or terms <b>will not be retrieved</b> (narrows search).
	(bipolar AND depression) AND NOT psychotic	Multiple ANDs, ORs, AND NOTs can be used in the same search statement; must use parentheses.
Double Quotes: Phrase searching	"bipolar depression"	Phrase searching instructs the search engine to return only pages where the <b>terms appear in the exact order</b> specified.
Truncation/Wildcards: Special Characters are Useful		
*, ?, #	"bipolar depression" OR "manic depression" AND NOT psycho* (searches for both psychosis or psychotic)	Instruct the search engine to return multiple forms of a term and/or words that have most of the term in common, except for a letter or two. <i>Use thoughtfully. In this case, not only will psychosis and psychotic be found, but so will psychological, etc.</i>
Case Sensitivity:	Most search engines treat lower case searches as universal, but will perform a case sensitive search if any letter is capitalized. For example, search for the term web, the engine will return pages containing web or Web. Search for Web, only pages containing Web will be returned.	
Specify Document Type	To continue building on the examples used above, a search for clinical trials of medications used to treat manic-depressive illness, use the phrase "clinical trial" in the search: Ex., ("manic depressive illness" OR "bipolar depression" OR "manic depression") AND "clinical trials." Using trial* returns pages that mention either trial or trials.	
Utilize Expertise	If the purview of a particular query falls within the scope of a specialized search engine, use it. For example, if the goal of a search is images of medical instruments, then by all means, use a image search engine, such as Ditto. In addition, some centralized engines, such as Google, allow the search to be limited to images (Google Image Search). If, for example, the question involves federal law, then AllLaw is the place to go.	

## CONCLUSION

Time is well-spent in learning to pose a proper query and to take advantage of the way search engines work.

Search engines mentioned in the review:

Google: <http://www.google.com>

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

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

Ditto: <http://www.ditto.com>

AllLaw: <http://www.alllaw.com>

MedHunt: <http://www.hon.ch/MedHunt>