Internet Research: Sources and Strategies

**Major Concepts**

This section is an overview of Internet guides and search engines.

**Other Resources**

More about this topic can be found at:

- [http://tilt-web.at.nwu.edu/orient/research/](http://tilt-web.at.nwu.edu/orient/research/)

Beyond NU: Internet Searching Tools

- [http://www.library.nwu.edu/resources/internet/search/](http://www.library.nwu.edu/resources/internet/search/)

Search Engine Watch

- [http://searchenginewatch.com/](http://searchenginewatch.com/)
Getting the most out of Guides and Search Engines

Periodical databases, library catalogs, recipe files, Web pages... regardless of the format, it is often difficult to find the right piece of information or resource when you need it the most. Fortunately, there are some useful guides and search engines for finding and retrieving those needle-like resources in the huge electronic haystack we call the Internet.

Guides to information resources on the Internet

Try thinking of a Guide as a table of contents or a bibliography. With a Guide, you search for titles, subject matter, or categories. Organization schemes vary from guide to guide. Yahoo has one of the better subject classification schemes but minimal annotation. Guides that attempt to offer more annotation, greater selectivity, and even ratings information can be found in sites such as Point Communications, and Magellan. After searching Yahoo for a particular group or issue, you can often find similar resources in the same Yahoo subject category. If Yahoo cannot find anything in its own directory, it forwards your query to the Alta Vista Search Engine.

Search Engines

Search tools or search engines attempt to retrieve resources (usually Web pages) containing the terms that describe your topic. A software program (sometimes called a spider, robot, or webcrawler) explores and scans the Internet, gathering information about the location and content of various Internet sites. It then uses this information to create a searchable index for people to use. Try thinking of a Search Engine as somewhat analogous to the index you find in the back of the book that tells you the pages where certain words appear. On the Web, you use this index or search Engine to search for occurrences of a specific word or phrase.
Search Engine Features

Search Engines vary in size, precision, level of indexing, speed, display options, ease of use, and search interface.

Most return a list of results ranked in decreasing order of probable relevancy. The search software does not really understand English or other human languages — it just matches words.

The following section describes some common search engine features. Almost all Search Engines provide search tips or help documentation that you must read in order to determine the availability of these and other features.

Looking for more than one term (The Boolean AND)

Connecting search terms with AND tells the search engine to retrieve records containing all the terms. Doing a search with the connector AND or its equivalent will narrow your search. For example:

```
  oil AND embargo
```

limits the search to only those records containing both oil and embargo. As with many features, the implementation of Boolean searching varies from one search engine to another. Although AND may sometimes be the default option, it is also common to either select it as "all of these words" from a menu or type it directly into a search query. Another variation on AND is a + (plus sign) placed immediately in front of a desired term to indicate that it must be present in your search results.
Looking for any term (The Boolean OR)

The default search on many search engines is OR. Some Search Engines make an equivalent to OR available as an “any of the words” option on a menu. Linking search terms with OR retrieves records that contain any or all of the terms you enter. If you are seeking information on a relatively obscure, hard to find subject, or a topic that has several different names, doing a search for synonyms with the connector OR will broaden your search. For example:

Geronimo OR Goyathlay
Eskimo OR Inuit

allows us to cover more than one possible name.

Looking for a term, eliminating another (Using the Boolean NOT)

The use of NOT or its equivalent tells a Search Engine to retrieve records containing one term but not another. For example:

Saturn NOT planet

is one way to focus a search on Saturn, the car manufacturer as opposed to Saturn, the planet. This technique is far from foolproof. An astronomy Web page discussing Saturn that does not actually mention the word “planet” can still be retrieved with the above search. Equivalents to NOT available on some Search Engines include a - (minus sign) as well as AND NOT placed immediately in front of the unwanted term. For example, in Alta Vista’s Advanced Search mode, you could type

Jordan AND NOT Michael

(NOTE: Alta Vista’s Advanced Query returns results in no particular order unless desired terms are specified in the Results Ranking Criteria box.)
Capitalization

Case-sensitive Search Engines can be useful when capitalization affects meaning. For example, when capitalized, the meaning of dull knife changes from a useless kitchen implement to the name of the 19th Century Northern Cheyenne leader Dull Knife. Case-sensitive searches can also make it easier to distinguish between words and acronyms such as “now” and “NOW.” Entering a search in lower case generally makes a search engine look for both upper and lower case occurrences. Using uppercase letters, where appropriate, typically retrieves a smaller set of documents since only exact matches can be made. As always, check the search engine “help” or “tips” documentation to determine the availability of this and any other desired search feature.

Phrase searching

Now widely available, phrase searching allows you to search for documents containing a group of words appearing together in a specified order. Either enclose the phrase in quotation marks or check for a phrase search option on the search engine menu. Although some Search Engines have a name search feature, phrase searching is useful for that function as well. The following can be searched as phrases:

“thousand points of light”
“Joyce Carol Oates”
“white collar crime”
“Bureau of Indian Affairs”
“you can’t roller skate in a buffalo herd”

Proximity

A retrieved Web page may contain your search terms and still be unrelated to what you hoped to find. A document with your search terms in close proximity to one another is more likely to be related to your topic than a document in which those terms appear far apart. Although relevancy ranking should, in theory, give a greater “score” to the former document, you may need proximity operators to achieve greater precision.

quilt ADJ patterns

Tells search engines such as Webcrawler and Lycos to return results that have the words quilt and patterns next to one another.
Truncation

Many search engines allow you to take advantage of truncation or wildcard features when you want to make sure you have included all possible endings for a word or need make allowances for possible variations in spelling. For example, truncating memory to memor* tells Alta Vista to retrieve memory, memories, memorize, memorization, etc. Not all Search Engines use the asterisk * as their truncation or wildcard character. Check the help documentation to determine whether this feature is available in the Search Engine you are using.
<table>
<thead>
<tr>
<th>And</th>
<th>Or</th>
<th>Not</th>
<th>Phrase Searching</th>
<th>Proximity</th>
<th>Truncation</th>
<th>Capitalization</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Excite</strong></td>
<td>Power Search: Select &quot;MUST&quot; contain option. (Do not use +.)</td>
<td>Power Search: Select &quot;CAN&quot; contain option.</td>
<td>Power Search: Select &quot;MUST NOT&quot; contain option. (Do not use -.)</td>
<td>Not available.</td>
<td>Not available.</td>
<td>Not available.</td>
</tr>
<tr>
<td><strong>HotBot</strong></td>
<td>Use the HotBot drop-down menu. Select the &quot;all the words&quot; option (usually the default search).</td>
<td>Use the HotBot drop-down menu. Select the &quot;any of these words&quot; option.</td>
<td>Use the HotBot drop-down menu. Select &quot;open all&quot; for extra options. Change &quot;must&quot; contain to &quot;must not&quot; contain, then type the undesired term in the adjacent search box.</td>
<td>Not available.</td>
<td>Not available.</td>
<td>Available only for unusual types of capitalization.</td>
</tr>
<tr>
<td><strong>Infoseek</strong></td>
<td>Use + (plus) in front of search terms that must appear in the results.</td>
<td>OR is the default.</td>
<td>Use - (minus).</td>
<td>Enclose phrase in quotes.</td>
<td>Not available.</td>
<td>Yes.</td>
</tr>
<tr>
<td><strong>Lycos</strong></td>
<td>Regular Search: Not available. Use Lycos Pro or the Java Power Panel.</td>
<td>Lycos Pro: Use Lycos Pro drop-down menu. Select the &quot;Any of the words&quot; option.</td>
<td>Lycos Pro: Use Lycos Pro drop-down menu. Select the &quot;Boolean Expression&quot; option and use OR between search terms.</td>
<td>Regular Search &amp; Lycos Pro: Enclose phrase in quotes.</td>
<td>ADJ</td>
<td>Automatic. ($ was the truncation symbol but it no longer works very well.)</td>
</tr>
<tr>
<td><strong>Webcrawler</strong></td>
<td>Use AND between search terms.</td>
<td>OR is the default search. No need to specify.</td>
<td>Use NOT in front of the undesired search term.</td>
<td>Enclose phrase in quotes.</td>
<td>NEAR# (as in NEAR/25) ADJ</td>
<td>Not available.</td>
</tr>
<tr>
<td><strong>Yahoo</strong></td>
<td>Select &quot;Matches on all words (AND)&quot; from search options menu or use + (plus).</td>
<td>Select &quot;Matches on any word (OR)&quot; from search options menu.</td>
<td>Select &quot;an exact phrase match&quot; from search options menu or enclose phrase in quotes.</td>
<td>Not available.</td>
<td>Yes. Use *.</td>
<td>Not available.</td>
</tr>
</tbody>
</table>