

# Internet Search Techniques

Tariq Ghayyur



# Web Search Basics

At the end of 2001 the web contained over 2 billion pages and Google searched 500.000 pages. In Jan of 2005 Google searched more then 8 billion pages. Even more impressive is the fact that it can be relatively quick and easy to find information on those pages.

Note: many search engines may contain multiple methods e.g. Yahoo at one time used Google's search engine for some of its searches.



# Some Basic Terminology

---

- Website: Collection of web pages
- Search Engine: Search documents for specified keywords
- Web Page: hypertext document on the World Wide Web that can incorporate text, graphics, sounds
- Web Browser: An application that opens websites
- Internet Protocol: Set of rules that govern over internet

# Search Techniques

---

Search techniques are the methods you use to search catalogs, article databases, and web search engines to find information.

The use of the Boolean operators AND and OR between keywords to create a search statement is an example of a search technique.

# Quotation Marks

---

Quotation marks are used around phrases. By using quotations marks, you are telling the computer to only bring back pages with the terms you typed in the exact order you typed them.

Example:

“health care reform”

*instead of*

health AND care AND reform

# Truncation

---



Truncation means to chop off. When you truncate you chop off the end of the word, so the computer can search for multiple endings.

For example, your research question includes the keyword **education**. You can truncate education, so that the computer will find all of the word ending variations.

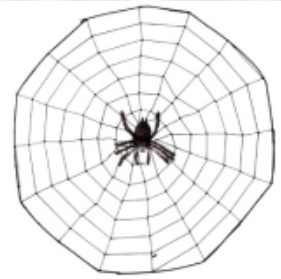
Educat\* will find:

Education

Educate

Educated

Educating



# Web Searching Techniques

---

- Use quotation marks to search for phrases
- Use the minus (-) symbol in front of a word to exclude it
- Use the advanced search screen to help refine your search
- Be very specific
- Search in appropriate domains (.com, .edu, .gov)

# Domain Searching

---

Finally, when searching the web, it is often helpful to limit your results to a specific domain.

The domain of a web site indicates the type of organization or geographic region from which the site originates. Helpful domain limiters include:

- .edu-education
- .gov-government,
- .mil-U.S. military,
- .org-organization
- .com-commercial



# Search strategies

---

- Do not
  - use word “search”
  - use a string of keywords without specifying Boolean properties
  - use upper case unless part of strategy
  - use NOT or - unless absolutely sure is necessary
    - **elimination of unanticipated pages**

# Types of search engines

---

- Keyword or robot based (builds a database)
- Directory based (categories indexed by people rather than computer)
- Meta indexes (can combine searches or allow you to search a variety of engines individually)
- Specialized search engines

# Keyword or robot based Search Engines

---

- Large database of web pages
- No human involvement and no quality control
- Can submit website or will find some on own
- Searches full text to certain level, does not search deep or invisible web
- Google ([www.google.com](http://www.google.com))
- Alta Vista ([www.altavista.com](http://www.altavista.com))
- Fast ([www.alltheweb.com](http://www.alltheweb.com))
- Wisenut ([www.wisenut.com](http://www.wisenut.com))

# Directory-based Search Engines

---

- Indexed by individuals so subject searches will be more accurate
- Smaller database than Robot engines
- Used mainly for finding good site on general topic
- Yahoo ([www.yahoo.com](http://www.yahoo.com) )
- About ([about.com](http://about.com) )
- Looksmart ([www.looksmart.com](http://www.looksmart.com))

# Meta indexes

---

**Meta-Searches**; these organize links received from searches by more than one search engine.

- ❑ <http://www.metacrawler.com/>
- ❑ <http://www.ixquick.com/>
- ❑ <http://www.dogpile.com/>
- ❑ [www.surfwax.com](http://www.surfwax.com)
- ❑ <http://www.clusty.com>

# Specialized Search Engines

---

- Geographic based ([www.altavistacanada.com](http://www.altavistacanada.com))
- Phone directories ([canada411.sympatico.ca](http://canada411.sympatico.ca))  
Phone directories ([canada411.sympatico.ca/](http://canada411.sympatico.ca/), [www.infospace.com/canada/index.htm](http://www.infospace.com/canada/index.htm))
- Newsgroup searching ([groups.google.com](http://groups.google.com))
- News searching ([news.google.ca](http://news.google.ca))
- Women's information ([wwwwomen.com](http://wwwwomen.com))
- Different formats ([www.gimpsy.com/Different](http://www.gimpsy.com/Different)  
formats ([www.gimpsy.com/](http://www.gimpsy.com/), [www.kartoo.com/](http://www.kartoo.com/)  
)

# To find more info on search engines

---

- Searchenginewatch ([www.searchenginewatch.com](http://www.searchenginewatch.com))
- Searchengineshowdown ([www.searchengineshowdown.com](http://www.searchengineshowdown.com))

