Chapter 5:
Find What You Want - Fast

Web 101
Third Edition

by Wendy G. Lehnert &
Richard L. Kopec
Learning Objectives

- Find out how to analyze your information needs in order to select appropriate tools for the job
- Learn how to search subject trees and clearinghouses for useful information and resources
- Discover how to use successive query refinement when you visit general search engine
Learning Objectives

• Explain how and when to select a new query mode
• Find out about advanced search features and specialized search engines
• Find out how to assess the credibility of information of the Web
Taking Charge

- Search engines and meta search engines have become powerful.
- People can use search engines without knowing much about using them or how they work.
- Search engines are designed to be easy to use, but you still need a strategy when looking for information.
Taking Charge

• Finding information quickly is important
• Analyzing your information will allow you to select the right resources
• Selecting appropriate search queries will allow you to begin the process of finding exactly the information that you are looking for
When you begin the search process, you should decide which type of question you need answered:

- Voyager question: an open-ended exploratory question
- Deep thought question: a more focused open-ended question with a specific goal that may have multiple answers
- Joe Friday question: a very specific question that has a simple answer
Taking Charge

- Voyager and Deep Thought questions require input from multiple documents
- Browsing is part of the exploratory process
- Joe Friday require facts, which will likely come from one document
Taking Charge

• The next step is to select the type(s) of Web resources needed to help you find your answers
  – Subject tree
  – Clearinghouse
  – General search engine
  – Specialized search engine

• You may need to ask more questions
Taking Charge

• A Subject tree is
  – A hierarchy that is organized by topics
  – Each topic has an associated list of Web sites and online documents
  – You navigate through the hierarchy to find information
  – Also called directories

• A Clearinghouse is a collection of Web pages and online documents about a topic; may contain a hierarchy too
Taking Charge

- A general search engine is a search engine that indexes a large collection of Web pages via keywords.
- A Specialized search engine is similar to a general search engine, except that it is limited to a specific topic.
More About Subject Trees and Clearinghouses

• A subject tree is a browsing aid
• A subject tree has a tree-like structure
  – You start from the root of the tree
  – Then you branch out to topics, and you chose a new subtopic with each choice

• Examples include:
  – Yahoo (oldest and most popular)
  – About
  – Open Directory Project
More About Subject Trees and Clearinghouses

- Since subject tree’s have a specific organization to maintain, each new document is added by people
  - The content is checked
  - The appropriateness of the document insertion is checked
  - This does not guarantee the quality of information
- Organizational problems can still occur
- The structure is not always intuitive
More About Subject Trees and Clearinghouses
More About Subject Trees and Clearinghouses

- Some subject trees also have a search engine to help you search the site.
- The URLs that you see listed in a subject tree are not for pages located on that site; they will be on another site.
More About Subject Trees and Clearinghouses
More About Subject Trees and Clearinghouses

• When trying to answer an Deep Thought question using a subject tree:
  – Look for relevant topics that relate to the question
  – Think of keywords to help you during browsing or using the search engine
  – Keyword creation may require associative thinking

• Also, try considering the larger context in which the original question was posed
More About Subject Trees and Clearinghouses

Copyright © 2007 Pearson Education, Inc. Publishing as Pearson Addison-Wesley
General Search Engines and Meta Search Engines

• Before using a general search engine, you will want to keep some points in mind
  – Know your search engine, as the structure of the query can vary
  – Never look beyond the first 20 to 30 hits for any given query
  – Experiment with different keywords in different queries
  – Don’t expect the first query you try to be your last
General Search Engines and Meta Search Engines

• For a Joe Friday question, the search engine Ask.com is a good place to start
  – Type in your question in the form of a question (Who invented the television?)
  – Ask.com returns a list of questions for which it has answers for
    • Some may be what you’re looking for
    • Other may not be related at all
  – The selected question will return a list of answers
General Search Engines and Meta Search Engines

who invented the telescope - Ask.com Web Search - Mozilla Firefox

Web Search

Compare Prices at BizRate
Find Bargain Prices on Who Invented The Telescope!
www.BizRate.com

who invented the telescope [Web Answer]
Allah's problem lost importance after Galileo Galilei invented the telescope. Unfortunately there is no opportunity here to discuss this fundamental change from mirrors to lenses and later on to Newton's prism.

Who Invented the Telescope?
I know that Galileo was a famous astronomer, but I don't think he invented the telescope.
www.surveycentral.org/survey/13121.html • Save

Extraterrestrials, Mars meteorite - Books - The Washington Times
... nature of the moon using the recently-invented telescope opened speculation about lunar life by distinguished figures including Johannes...
www.washtimes.com/books/20060725-104347-8786r.htm • Save

The First Telescope
The excited children show their discovery to their father, and thus the telescope is invented. News spreads across Europe to Galileo, and the...

Narrow Your Search
Invented the First Telescope
Who Invented the Microscope
Invention of the Telescope
History of the Telescope
Who Invented the Computer
Who Invented the Telephone
Who Invented the Radio
Expand Your Search
How Does a Telescope Work?
What Does Equinox Mean
Why Does It Rain
How Long Is a Year on the Planet Mercury?
Who Discovered Pluto
Related Names
General Search Engines and Meta Search Engines

- If you can’t find what you are looking for using one search engine,
  - You may need to try another one
  - You can also try a different query
- A Meta Search Engine allows you to access multiple search engines
  - Saves you time
  - You use it to send a query to multiple search engines; best with simple queries
General Search Engines and Meta Search Engines

• Some engines provide clustering, for easy inclusion or removal of categories in the results
Tools of the Trade

• Don’t expect to find the best hits on your first try
• Each new search problem will give you an opportunity to refine your queries until you find what you are looking for
• Query refinement may consist of
  – Narrowing if you get too many hits
  – Broadening if you get only a few good enough hits
  – Redirecting if nothing useful comes up
Tools of the Trade

- Several search tools can assist you with query refinement
- Exact phrase matching allows you to enter a phrase that is treated as a single keyword
  - Documents that contain this phrase are ranked higher than if the words were in separated
  - For example: “chocolate cream pie”
  - Used for narrowing or redirecting a search;
Tools of the Trade

• NOT is used to redirect a query
  – If you find that a keyword or phrase is connected to irrelevant hits, then add it to your query as AND NOT keyword
  – For example, cookies AND NOT JavaScript

• You can combine keywords and operators to create complex queries
  – Use parenthesis () for grouping
  – For example, X AND (Y OR Z)
Tools of the Trade

• Wildcards are useful for retrieving variations of a word
• For example, art* will search for art, artwork, artist, artistry, and so forth
• An excellent way to broaden a search
• Different wildcard characters are used by different search engines
• The most common characters are: *, #, and ?
Tools of the Trade

• If you are looking for something specific, then a Boolean query is useful.

• A Boolean query consists of keywords combined with the Boolean operators AND, OR, NOT.
  – X AND Y returns pages with both X and Y.
  – X OR Y returns pages with X, Y or both.
  – X AND NOT Y returns pages with X that do not contain Y.
Tools of the Trade

- Boolean queries are flexible enough to handle all three types of query refinement
- Boolean queries correspond to formal logic
- Some search engines have a special Boolean search mode, such as AltaVista
- Search engines are not forgiving if you make a mistake in your query
Tools of the Trade

- AND is used for narrowing a query
- If you know that your target documents will contain a group of keywords, list them using the AND operator
- You can combine exact phrases with the OR operator too
- OR is used for broadening a query
- If you can think of related words for a topic, list them using the OR operator
Tools of the Trade

• NOT is used to redirect a query
  – If you find that a keyword or phrase is leading irrelevant hits, then represent it in your query as AND NOT keyword
  – For example, cookies AND NOT JavaScript

• You can combine keywords and operators to create complex queries
  – Use parenthesis () for grouping
  – For example, XAND (Y OR Z)
Tools of the Trade

[Image of an advanced web search window with options for querying search terms, Boolean operators, and date range.]
Tools of the Trade

• You can use required and prohibited keywords to filter documents with certain words
  – + means that the word is required
  – For example, +apples
  – - means that the word is prohibited
  – For example, -apples
• Useful during query refinement
• You can combine this with phrases too
Searching the Invisible Web

• Many Web pages are created “on the fly” using content extracted from a database in response to Web page query.

• Spiders (aka Web crawlers) that gather information about Web pages for the search engine cannot access these databases.

• Consequently, this information can’t appear in the search engine database - the Invisible Web.
Searching the Invisible Web

• Try using a general Web directory
  – Librarians Index (http://lii.org/)
  – AcademicInfo (http://www.academicinfo.net/)
  – InfoMine (http://infomine.ucr.edu/)

• Can also use Google or Yahoo!, but add the word *database* to your keyword list
Searching the Invisible Web

Topographical Bibliography
Topographical Bibliography of Ancient Egyptian Hieroglyphic Texts, Statues, Reliefs, ...
These list ancient Egyptian monuments still in situ, those found in ...
www.ashmolean.ox.ac.uk/gru/3.html - 4k - Cached - Similar pages

PA Piccione: Bibliographical Database of Ancient Egyptian Medicine
Classified Bibliographical Database of Ancient Egyptian Medicine and Medical Practice. List of Subject Headings. The numerals before each subject heading ...
www.cofc.edu/~piccione/medbase.html - 39k - Cached - Similar pages

Making up of an Astro-Egyptology related database
The idea of creating an astro-egyptological database comes from the ... This dossier is not a general introduction to ancient Egyptian Astronomy ...
www.culturedf.org/english/astroegypt1.htm - 24k - Cached - Similar pages

Open Directory - Society: History: By Time Period: Ancient: Africa ...
Ancient Egypt by History Link 101 - Links to sites of art, biographies, ... presents 5000 years of Egyptian history and culture using a database of ...
dmoz.org/Society:History/By_Time_Period/Ancient/Africa/Egypt/ - 31k - Cached - Similar pages

Ancient Egypt goes online with IBM: WebWatch - Breaking Business ...
The Eternal Egypt project has created an online digital database of Egyptian ...
This site has time visits how ancient Egyptian functions such as the ...

Copyright © 2007 Pearson Education, Inc. Publishing as Pearson Addison-Wesley
Searching the Invisible Web

• Most browsers now include an integrated search toolbar
• IE v7 also allows the user to add several search engines to the list
Searching the Invisible Web
Assessing Credibility on the Web

• You need to have a critical eye when examining the Web for legitimate research.
• The Web contains both good and bad information.
• Focus on the information rather than the page’s appearance.
• Many factors should influence your decision to use information found online.
Assessing Credibility on the Web

• The following are some common criteria to keep in mind when evaluating content
  – Author credibility
  – Accurate Writing and Documentation
  – Objectivity
  – Stability of Web pages
  – Fraudulent Web pages
Assessing Credibility on the Web

• Author credibility
  – Information is useless if the author is not identified
  – The author’s contact information should also be available
  – Try to verify that the author is who he/she claims to be
  – Find out if the author is qualified to write on that topic
Assessing Credibility on the Web

• Accurate writing and documentation
  – Sloppy writing can mean that the content is of poor quality
  – Check out any references to other work
  – No references may be a red flag

• Objectivity
  – If the author is associated with a company, separate the content from advertising
  – Take sources of funding into consideration
Assessing Credibility on the Web

• Stability of Web pages
  – Information comes and goes
  – Does the page have a date?
  – When was the page last revised?
  – The life of the average page is 75 days

• Fraudulent Web pages
  – It is easy to create a Web page in someone else’s name
  – Some of these pages are parodies, but others may be malicious