

Successful Web Search Strategies



Kathleen Schrock

Statistics from www.emarketer.com

- There are about 72 million domain names on the Net
- There are about 2.1 billion Web pages
- Any one search engine can only provide feedback on 40% of Web information



Internet Software Consortium, 2000

Itinerary

- Definition of “search engine”
- How a search engine works
- Definition of “subject directory”
- Improving your use of search engines
- Effective search strategies
- Teaching students to search
- The future of searching



What is an Internet search engine?

- An Internet tool which will search for sites containing the words that you designate as a search term
- Search engines search their own databases of information

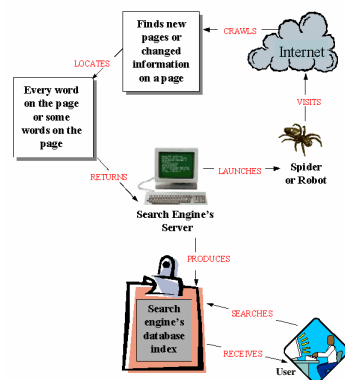


How does a search engine work?

- Deploys a robot program called a *spider* or *robot* designed to track down Web pages, follow the links these pages contain, and add information to their own database
- Each search engine has its own way of doing things



How does a search engine work?



How do the search engines differ?

- Different databases are searched
- The degree of detail collected differs
- The level of sophistication of the “robot” varies
- Relevancy rankings differ



What determines relevancy?

- Results are presented in order of relevance
- Web page designers often try *spamdexing*
- Relevancy rankings based on a formula
- Differing methods of collection

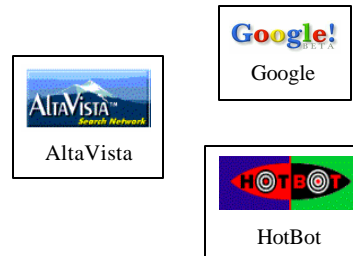


How does a search engine know about the documents on the Web?

- If the search engine’s “spiders” or “robots” find the site on their gathering missions
- If the publisher of the document registers it with the search engines



What are some search engines?

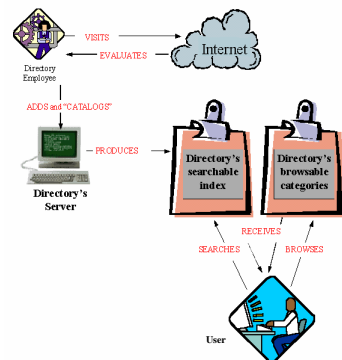


What is an Internet subject directory?

- Organizes Internet sites by subject
- User works down a series of menus
- Smaller than search engines
- Maintained by human beings

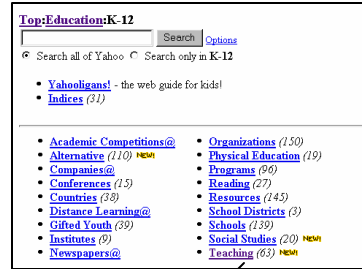


How does a directory work?





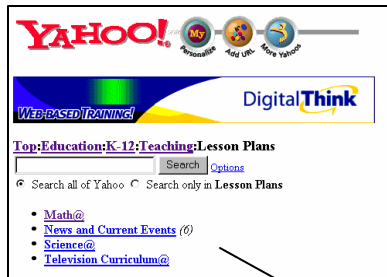
Choose "Education" from the top menu



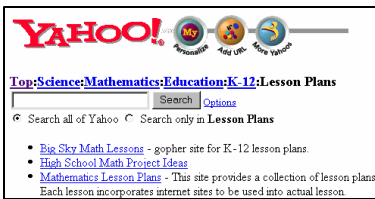
"Teaching" category



Two sub-categories



Lesson plans choices



Math lesson plans

Which is better -- a search engine or a subject directory?

- Personal preference
- Directories allow the user to form the search, but may not use common terms
- Search engines allow the user to use common language, but may lead to lots of hits



How do I prepare to search?

Formulate your search question carefully

Examples

- I want to find information on popular methods of losing weight.
- What are the characteristics to look for when purchasing a diamond?
- When do the Monarch butterflies migrate?



How do I prepare to search?

Identify keywords

I want to find information on popular methods of losing weight.

What are the characteristics to look for when purchasing a diamond?

When do the Monarch butterflies migrate?



How do I prepare to search?

Select synonyms and other word forms

I want to find information on popular methods of losing weight.

Popular _____ common favorite

methods _____ method ways way techniques technique

losing _____ lose reducing reduce reduction

weight _____ fat dieting diet



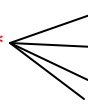
How can I improve my use of search engines?

- Read the help files
- Learn about wildcard searches
- Learn about Boolean searching



Wildcard Searches

- A special symbol which allows you to search simultaneously for several words with the same stem

educat*  educator
educators
education
educational



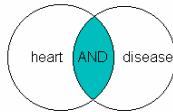
Boolean Searches

- Allow you to broaden a search
- Allow you to narrow a search
- Designated differently in each search engine



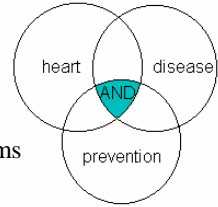
How do I narrow a search using “and”?

- Using *and* will only provide links to sites which have BOTH of these words present



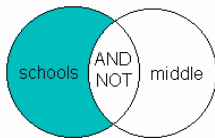
How do I narrow a search even more?

- Using *and* twice will limit the search even more
- You will only get pages that include all THREE of the terms



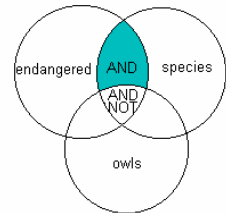
How do I narrow a search using “not”?

- Using *not* narrows the search by telling the search engine to exclude certain words
- AltaVista uses the form *and not*



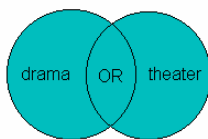
Can I combine two different operators?

- You can combine different operators to limit or broaden a search



How do I broaden a search?

- Use of the word *or* will broaden a search
- Use *or* if two words may be used interchangeably
- Can use *or* more than once to get very broad

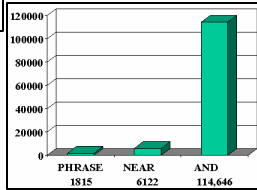


How do I narrow a search using “near”?

- Using *near* narrows the search by telling the search engine to look for words close to others
- Use when *and* gives too many hits

“Near” comparison

“prevent heart disease”
“heart disease” near prevent
heart and disease and prevent



How do search engines deal with Boolean operators?

- Not consistent among search engines
- May enter it by hand or via a pop-up box
- May have to go to an “advanced” page
- The default may be *and* or *or*



How do I prepare to search?

Combine synonyms and keywords carefully

Combine synonyms with Boolean OR
(popular or common or favorite)
(losing or lose or reducing or reduce or reduction)

Use wildcards to combine variant word forms
(method* or way* or technique*)
(weight or fat or diet*)

Combine these phrases with the Boolean AND to limit the search

Dieting methods search string

Combine synonyms and keywords carefully

(popular or common or favorite)
and (method* or way* or technique*)
and (los* or reduc*)
and (weight or fat or diet*)



Result #1 from HotBot

WEB RESULTS 228,500 Matches

1. Alpha Medical Nutrition: Self-Management Solutions For Common Health Problems
an educational resource with 20 Medical Centers that explain common diseases and their solution. Self-care and prevention are emphasized through nutrition, exercise and improvements to the environment. Support with course consultations, programs & 25/2000 <http://www.alphaexpress.com/>
See results from <http://www.hotbot.com/>

Some other considerations

- Know how to use your browser
- Consider the time of day
- Check your spelling
- Search engines are in competition



Excite : Power Search

What is a “meta” search engine?

- A meta search engine doesn't create its own database of information
- It searches the databases of other engines
- Allow you to sift and see what is out there



What is a unified search interface?

- Users enter search queries to several search engines/directories within the same interface
- The databases are searched sequentially, rather than simultaneously



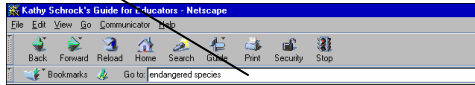
All-In-One Search Page http://allonesearch.com/

Specialized Interest from All-In-One

Searching shortcuts in Netscape

Netscape Navigator 4.0/Communicator

Type your keywords/search phrase into the Location window and hit enter



Searching shortcuts in Internet Explorer

Internet Explorer 5



Using the frame on the left which appears when you choose *Search*, you can view the results of your search in the right-hand frame.

Choosing the right terms

- Use a unique keyword
- Make it a multi-step process
- Search only your results
- Start with a “title” search
- Find out if case counts
- Check your spelling
- Try using quotation marks around phrases



What are some effective search strategies?

- Decide whether you need a subject directory or search engine
- Be as specific as possible when using a search engine
- Try different search engines
- Read the help and tips files for each search engine



How do I determine if a search engine meets my needs?

- How easy was it to figure out how to search?
- Were the help files adequate?
- Were the amount of results returned adequate/overwhelming?
- How current were the results?
- What else makes it good?



What is the best tool for the job?

What is the best tool for the job?

INFORMATION NEED

- I have an idea of a **broad**topic.



Broad topic : Yahoo

The screenshot shows the Yahoo! homepage with the following elements:

- Navigation icons: Yahoo!, Mail, News, Finance, Sports, Weather, Local, More.
- Search bar with a "Search" button and a "Options" link.
- Links for "World Series Game 6", "College Football", "NFL", and "NBA Preseason".
- Links for "Yahoo! Pages", "Simple Search", "Maps", "Classifieds", "Personal", "Chat", "Email", and "My Yahoo!".
- Four columns of category links: Arts and Media (Arts), Business and Economy (Business), Computers and Internet (Internet), Education (Education), News and Media (News), Recreation and Sports (Sports), Reference (Reference), and Regional (Regional).

What is the best tool for the job?

INFORMATION NEED

- I want a small number of **relevant hits** and an idea of what's in each document before I visit the page.



Small number of hits & good summaries: Excite

The screenshot shows search results from Excite for the query "Digital Art Galleries".

- Header: "Top 10 matches" and "Show Titles only List by Web site".
- Summary: "71% The Fractal Microscope".
- URL: http://www.ncsa.uiuc.edu/Edu/Fractal/Fractal_Home.html
- Summary: "The Fractal Microscope is an interactive tool designed by the Education Group at the National Center for Supercomputing Applications (NCSA) for exploring the Mandelbrot set and other fractal patterns." "Robert M. Fanooff working with the Education Group at NCSA, it is possible to explore many common elementary mathematical principles while examining the Mandelbrot set."
- More Like This: [Click here to perform a search for documents like this one.](#)
- Summary: "67% Fractal pictures & animations".
- URL: <http://www.cnam.fr/fractal.html>
- Summary: "Another one (more recent) exists here. Check out this 2D & 3D Mandelbrot fractal explorer (set up by Robert Keller). International access for this archive - America continent. A problem has occurred on the server hosting this archive: only 2 by users can retrieve data (instead of 30 usually)."
- More Like This: [Click here to perform a search for documents like this one.](#)

What is the best tool for the job?

INFORMATION NEED

- What's **available on the Internet** for my topic?



What is available for my topic : MetaCrawler

The screenshot shows search results from MetaCrawler for the query "fractals".

- Query: fractals
- Collated Results: 1 to 20 of 47 references.
- Results page: 1 - 1 - 2 next
- 1000 The Fractory: An Interactive Tool for Creating and Exploring Fractals
- Excite: We have separated the information on this page into the five levels listed below in increasing order of difficulty. If you are a programmer, try writing some programs to generate fractals using the equations from level 4. Post messages on our message boards if you have questions (or answers), and try some references from our Links and References page.
- Infoseek: An interactive tool for creating and exploring fractals.
- 847 Web Guide - Fractals and Fractal Art
- A gallery of images produced with Pascal's Attractors Co software. sci fractals FAQ. A lot of useful information gathered from the sci fractals newsgroup. [star] Carlson's Fractal Gallery. 3D Phoenix Spirals, Atan Method Fractals, 3D Stalks, Bubbles.
- 847 Fractals
- noframe version.

What is the best tool for the job?

INFORMATION NEED

- I need to do a **pinpoint search** because my terms are narrow.



Pinpoint search : AltaVista

About 10132 documents match your query.

- Fractals**
noFrame version.
http://www.dna.nl/~fredbear/fractals/fractal_page.html - size 346 bytes - 27-May-97
- RenderVox Engine: Fractals**
Fractals: Met fractalbestemmings en de techniek van computers is het mogelijke heel kunstzinnige plaatjes te maken, fractals genaamd: Een fractal is een...
<http://home.gysost.nl/~blauw/FRACTALS/Fractals.html> - size 3K - 30-Dec-96 - Dutch
- Infoseek: Fractals**
shop - Infoseek Home You are here Directory > Arts & Entertainment > Digital art > Fractals...
http://guide.infoseek.com/Arts_and_Entertainment/Digital_art/Fractals - size 8K - 17-Aug-97 - English
- The Topical Pages: Fractals**
Iterated Function Systems Made Easy. Again, there's not much here as of yet. I hope to create a layman's explanation of IFS's, perhaps even with a form...
<http://www.personal.umich.edu/~darrrenlo/fractals/index.html> - size 706 bytes - 17-Aug-97 - English

What is the best tool for the job?

INFORMATION NEED

- I have **common keywords** that probably appear in many documents and I should make my search specific.



Common keywords : HotBot

Returned: 35327 matches.

Breakdown: fractals 38839

1 - 10

- galena.tj.org
99% galena.tj.org galena is a 386 running Linux and holds the distinction of being the first webserver at Thomas Jefferson School. Wow. It also contains the Thomas Jefferson School student web pages and other oddities. Items on this server Thomas...
<http://galena.tj.edu/inter.net/>, 2334 bytes, 26Jan97
- [Dynamical Systems: Java: Fractals: Koch Curve](http://www.math.uakron.edu/~kramton/java/fractals/roch/curve)
99% DYNAMICAL SYSTEMS: JAVA: FRACTALS: KOCH CURVE
ACKNOWLEDGEMENT Applet by SUN Microsystems. COMPUTER SCIENCE
DISCRETE STRUCTURES DYNAMICAL SYSTEMS FORTRAN
MATHEMATICS NUMERICAL ANALYSIS WWW Maintained By: Pankaj
Kamthan , kamthan@cs.concordia.ca...
<http://indy.cs.concordia.ca/www/java/ds/koch-curve/CLSFractal/>, 2878 bytes,
26Sep97

What is the best tool for the job?

INFORMATION NEED

- I have common **keywords in a phrase** like *New* in New Orleans or *A* in Vitamin A that cannot be ignored.



Keywords in a phrase : Infoseek

Related topics <ul style="list-style-type: none">CancerHealth	Search results 1 - 10 Hide Summaries next 10
Related news Search for recent news about Vitamin A	[2116] Maternal Serum Vitamin A Levels Are Not Associated with 1997 NCWH Abstracts Archive. Select your destination: Main Page New on the IAFAC Web Annual Therapies Conference Consumer Information Minority Issues Miscellaneous Issues ... 73% http://www.spaac.org/women/serum/2116.html (Size 5.6K)
Infoseek classifieds Find anything from concert tickets to computers, pool tables to personals	Vitamin A May Fight Melanoma Desired to be healthy long, Thure takes an upbeat approach to some of the hottest lifestyle issues for adults today, including health, fitness, nutrition and sex. 73% http://www.thuronline.com/@/01/17/01/1A43E03p5/sharv/theinfoLibray/news/98/19978p2227.html (Size 10.8K) More results from this site...
	Combat - Immune Supporting Nutrients Premature aging of the skin due to excessive exposure to sunlight is an increasing problem across the world today. 72% http://www.galt.com.com/~saw/fox2.htm (Size 6.2K)

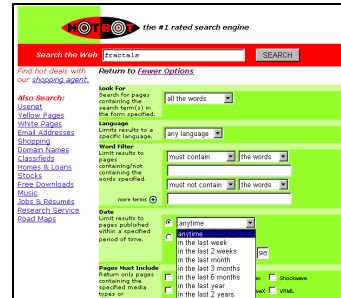
What is the best tool for the job?

INFORMATION NEED

- I know the **date of an event** and am looking for more information.



Date of an event : Hotbot



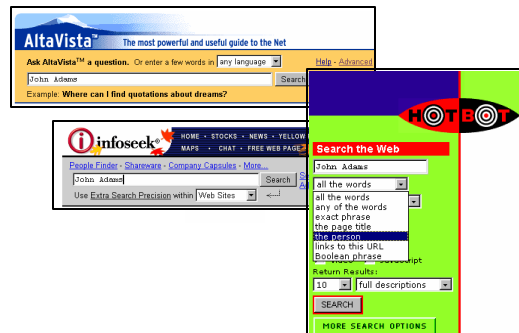
What is the best tool for the job?

INFORMATION NEED:

- I need information on a **proper noun**.



Proper noun : AltaVista, Infoseek, Hotbot



What is the best tool for the job?

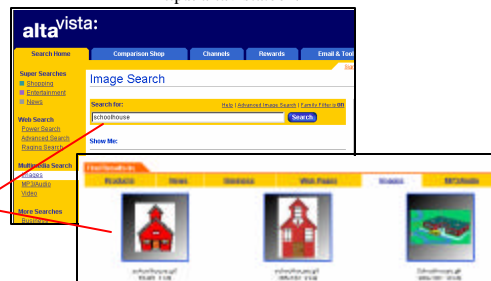
INFORMATION NEED

- I want to find an **image** on the Web.



Image : AltaVista

http://altavista.com/



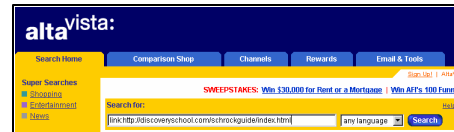
What is the best tool for the job?

INFORMATION NEED

- I want to find out who has **links to a site** .



Links to a site : AltaVista



Common error messages

- 404 Not found
 - host computer found, but not the file you are looking for
- 403 Forbidden/Access denied
 - Web site needs special permission or a password to access
- 503 Service unavailable
 - server is very busy; try again in 30 seconds
- Document contains no data
 - you are at the right place but the page is being updated at that moment; try again later

How do you teach students to search?

What would be a useful student search strategy?

- State their purpose in one sentence.
- List the databases they plan to search.
- Brainstorm keywords and topics.
- Consider varying word forms.
- Refine the search online
- Offline evaluation

Valenza , Joyce. *Power Tools*. ALA, 1998.

What is the future of searching?

- Mapping
- Collaborative filtering
- Client-side metasearching
- Personal agents



What is the future of searching?

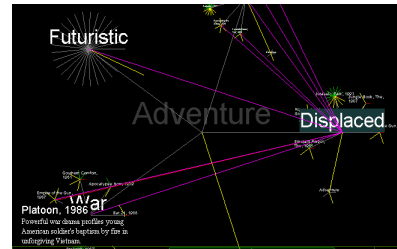
Mapping

Java-based products that analyze and organize documents by concept and then present the user with a dynamically generated navigable map of the relationships and hierarchies



Mapping : Perspecta

www.perspecta.com



What is the future of searching?

Collaborative Filtering

Using data on where other users who have visited a site have gone, as well as text analyses of the site, these engines dynamically suggest other links



Collaborative Filtering : Alexa's Internet

www.alexa.com

What is the future of searching?

Client-side Metasearching

A search tool that resides on your machine, searches the Net, cleans up the results, and returns the results to you.



Client-Side Metasearching : Copernic

www.copernic.com

Keywords	Words	Scheme	Updated	Matches	Category
endangered birds	All words	Quick search	4/15/99 16:03	57	The Web
<input type="checkbox"/>	U.S. Fish & Wildlife Service Home Page	http://www.fws.gov/			
<input type="checkbox"/>	Welcome to the U.S. Fish and Wildlife Service, a Bureau within the Department of the Interior. The mission of the U.S. Fish and Wildlife Service is working with al...	http://www.fws.gov/			
<input type="checkbox"/>	Keywords - Endangered Species.com	http://www.endangered-species.com			
<input type="checkbox"/>	We provide all the best endangered species information, endangered species books, and everything else regarding rare and endangered species, conservation...	http://www.massaudubon.org/br..._large.html			
<input type="checkbox"/>	MAC Bird Is Beyond Grassland Birds - Lang...	http://www.massaudubon.org/br..._large.html			
<input type="checkbox"/>	Mass Audubon's Audubon's leading efforts to restore a species list habitat to protect these rare birds.	http://www.far.com/bare/birds/area.html			
<input type="checkbox"/>	The World Wide Web Virtual Library - Animal...	http://www.wwvlib.org/animal.html			
<input type="checkbox"/>	This page has four subsections: General sites with focus on many types of animals; Specialized sites that focus on one type of animal; Miscellaneous documents a...	http://www.usgs.gov/			
<input type="checkbox"/>	Endangered Species - Birds	http://www.usgs.gov/			
<input type="checkbox"/>	Peregrine Falcon (Falco peregrinus) Southern Bald Eagle (Haliaeetus leucorhaphus)	http://www.usgs.gov/			
<input type="checkbox"/>	Books on Endangered Species: General - End...	http://www.endangered-species.com			
<input type="checkbox"/>	We provide all the best endangered species information, endangered species books...	http://www.endangered-species.com			
<input type="checkbox"/>	List of Endangered Species: Louisiana - End...	http://www.endangered-species.com			
<input type="checkbox"/>	We provide all the best endangered species information, endangered species books...	http://www.barbara-jordan.com/			
<input type="checkbox"/>	Nature Photography by Barbara Jordan	http://www.barbara-jordan.com/			

What is the future of searching?

Personal Agents

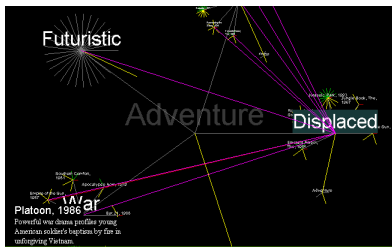
Users set up the agents with ongoing queries and the agents monitor the sites and notify the user of the relevant new pages



Personal Agents : Wisewire

www.wisewire.com

The screenshot shows the Wisewire website interface. At the top, there is a navigation bar with 'Home' and 'Help' links. Below that, the main heading is 'Geologic/Petroleum Engineering'. There are several tabs for navigation: 'New Finds', 'Best Of The Week', 'Best Of All Time', and 'Back to Home Oil'. A sub-heading says 'Sort by clicking on a heading below. Ratings are submitted by clicking a tab or a button below.' Below this, there are two main sections: 'Premium Content' and 'Web Pages'. The 'Premium Content' section has a table with columns for 'BUSINESS WIRE', 'Arakis Reports 1997 Nine Months Results', and a rating of '45 3'. The 'Web Pages' section has a table with columns for 'www.magstropix.com', 'Magnetic Pulse, Acoustic/Induction Borehole Measurements', and a rating of '71 4'. Another row in the 'Web Pages' section shows 'www.pls.com' and 'PLS Canada Inc. - Home Page' with a rating of '66 2'.



Kathy Schrock's Guide for Educators

<http://discovery.school.com/schrockguide/>



Based on
"Learning More about Search Engines
and Subject Directories"

http://www.cln.org/searching_faqs.html



"Sink or Swim : Internet Search Tools"

by Ross Tyner

[http://www.sci.ouc.bc.ca/libr/
connect96/search.htm](http://www.sci.ouc.bc.ca/libr/connect96/search.htm)



“Search Engines for the
World Wide Web”

by Alfred and Emily Glossbrenner

Peachpit Press, ©1998



“Searching the Net”

An Online Internet Institute Project

by

Janet Murray, Patty Sorensen, and Sheryl Steinke

<http://www.teleport.com/~janetm/oii/search.html>



“Choosing the Best Engine
for Your Purpose”

by

Debbie Abilock

[http://nuevaschool.org/~debbie/library/
research/adviceengine.html](http://nuevaschool.org/~debbie/library/research/adviceengine.html)



“Finding it Online:
Web Search Strategies”

by

Debbie Flanagan

<http://home.sprintmail.com/~debflanagan/main.html>

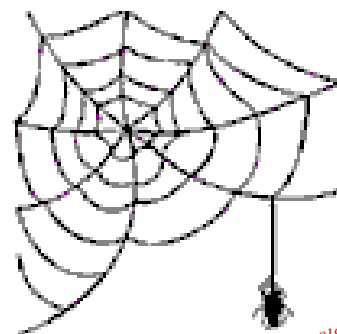


“The Future of Search”

by David Lidsky and Regina Kwon

PC Magazine, December 2, 1997

[http://www.zdnet.com/pcmag/
features/websearch/sb1.htm](http://www.zdnet.com/pcmag/features/websearch/sb1.htm)



The End

c1997-2002 Kathy Schrock