

Searching the World Wide Web

The world is indeed fascinated with the abundance of information that can be accessed from the World Wide Web. While the access to WWW information appears "instantaneous," if the researcher does not understand basic components and/or strategies associated with a WWW search, the information retrieved may be overwhelming, irrelevant, and unreliable. This tutorial, therefore, provides the "basic training" necessary for the completion of successful World Wide Web queries. Read and study each section of the tutorial at your own convenience and pace.

What is the Internet? What is the World Wide Web?

The Internet is a means of international communication in which information stored on any one computer (situated at any location around the globe) may be transmitted into the memory of another remote, distant computer. The information, therefore, may be considered found not "on" the Internet, but rather "through" it. Examples of communication service methods on the Internet include e-mail, the World Wide Web, Telnet and FTP (File Transfer Protocol).

The World Wide Web is one service component of the Internet which allows the individual operating the remote computer to access, through the use of a Browser (software which enables a transmittal of information, i.e. Netscape or Internet Explorer), documents, images, audio/video clips and voice transmittals. The determining factor regarding an adequate transmittal of the web page requested is the quality of the remote computer's hardware, as well as the associated software loaded thereon. Currently, approximately 3 billion documents may be accessed via the World Wide Web.

Web Sponsors







Web pages transmitted through the World Wide Web have specific sponsors, entities which hire individuals to create and post information pertaining to their purpose, cause or plan of action. Sponsorship of a web site is determined by the last three letters of a web "address" (also referred to as URL). An example of a web address is: <http://www.cccti.edu>. The last three letters (edu) indicate that the sponsor of the web site associated with this particular address is an educational organization. Varying sponsorship types include:


- .edu** Educational organization - information obtained is usually academically oriented. EX: <http://www.cccti.edu>
- .org** Non-profit organization - information presented to further a cause; economic profit is forbidden. EX: <http://www.samaritanspurse.org>
- .gov** Governmental association - information provided as a service of the local, state or national government - EX: <http://www.state.gov/>
- .com** Commercial business- information disseminated for the purpose of profit - EX: <http://eddiebauer.com>
- .mil** Military establishment - information specific to military purposes - EX: <http://www.af.mil>
- .net** Online community network - information prepared for a discussion group - EX: <http://www.thelearningsite.net>

NOTE: An understanding of sponsorship types helps the researcher to evaluate the authenticity and appropriateness of the information retrieved.

Accessing the World Wide Web

It is impossible for an individual working from a remote computer to access the totality of the seemingly infinitesimal number of World Wide Web pages. Instead, the researcher may find a World Wide Web address (webpage) specific to their particular topic via three distinct search tools detailed herein below:

WEB RESEARCH TOOL	CHARACTERISTICS	RECOMMENDED EXAMPLES AND THEIR FEATURES
Search Engines	<ol style="list-style-type: none"> 1. A search engine is a web tool containing multiple <u>keyword oriented</u> databases. (In other words, these databases index web pages by keyword, rather than subject). 2. Databases are compiled by computer-robot programs called "spiders" (hence the reference to the "World Wide WEB"). 3. A search engine scans the entire text of a particular web page in an attempt to retrieve the given keyword. Specific website addresses are therefore provided in the item retrieval list; browsing is not an option. 4. Many search engines earn revenue by placing advertisements or .com addresses at the beginning of a result list. Filtering is therefore essential. 5. Most search engines index only 20-30% of the web pages found on the World Wide Web, however, a few (such as Google) approach the 50% mark. 	<p> Currently, the largest, and perhaps the most popular, search engine. Indexes a variety of web page types (ex: pdf, Word, Excel, Power Point).</p> <p> SearchEdu.com retrieves web pages from academic organizations only! The information retrieved is therefore of a more scholarly nature.</p> <p> alltheweb.com allows the user to search by Media type (audio, video, picture, etc.), an excellent tool when preparing a multimedia presentation on a particular subject.</p> <p> Good search engine for current events.</p>
WEB DIRECTORIES	<ol style="list-style-type: none"> 1. Directories are tools which allow for hierarchical <u>subject category searches</u>. The websites contained in a certain category are hand-picked by editors, who are themselves experts in a specific field. 2. Full-text searches of the retrievable web pages do not exist; instead, each web page is placed within a search term "category." Thus, a retrieval <u>list</u> is broad, allowing the researcher to browse and find an appropriate sub-category. 	<p> dmoz is a well-organized, user friendly directory, all the while providing access to approximately 590,000 categories.</p> <p> Librarian's Index to the Internet provides access to categorized websites which provide "information you can trust," as our Country's librarians choose and maintain this directory.</p> <p>Academic Info categorizes information by academic discipline.</p>
INVISIBLE WEB	<ol style="list-style-type: none"> 1. The invisible web contains web pages that cannot be retrieved from search engines or directories. Many times this information contains searchable databases that are not stored on the visible web for economic or personal reasons. 2. Nonetheless, such databases are many times a valuable resource, rendering these often difficult searches worthwhile. 	<p>Using the recommended search engines, type in a keyword placing "database" behind it. EX: In a Google search, type "languages database" or "toxic chemicals database."</p> <p>Also, utilize the following search tools (though they are not user-friendly in the least):</p>

		 <p>A companion to the book <u>The Invisible Web: Finding Hidden Internet Resources Search Engines Can't See</u>.</p> <p>direct search Provides long page listings describing searchable databases on many academic topics. Compiled by Gary Price, an academic librarian.</p>
--	--	--

NOTE: The recommended sites described above may be accessed at the LRC'S Recommended Websites page: <http://www.cccti.edu/LRC/Extraneous/Internetaccess.htm>

Search Strategies

For each research tool discussed in the previous section of this tutorial, certain search strategies may be followed:

RESEARCH TOOL	STRATEGIES	OTHER TIPS
SEARCH ENGINE	<ol style="list-style-type: none"> 1. Use after you have narrowed a topic or for more specific searches. 2. To obtain information which will incorporate information on an entire search <u>phrase</u>, use quotations (EX: "childhood genetic disorders.") 3. Use succinct key terms or phrases. 4. Perform Boolean (use of and/or) searches (EX: Caesar Chavez and unions). 5. Search with alternate, synonymous terms. 	<ol style="list-style-type: none"> 1. Ignore commercial links provided in sidebars. 2. Do not use search engines for a general overview of a subject (use directories instead). 3. Complete the tutorials for your favorite search engine(s). 4. Avoid Meta-engines, such as "dogpile" or "lycos."
WEB DIRECTORIES	<ol style="list-style-type: none"> 1. Use for broad subjects, topic overviews and longer documents. 2. Utilize your browsing skills to find the most pertinent category for your search, as you will retrieve a list of sites relevant to your topic, rather than numerous individual web pages. 3. Use distinctive terms found in Subject Directories. 	<ol style="list-style-type: none"> 1. Avoid Boolean (and/or) searches as well as variant term searching. 2. Find a subject specific directory for your field of study (EX: http://www.ecwebguide.com for early childhood educators)
INVISIBLE (DEEP) WEB	<ol style="list-style-type: none"> 1. Use to access databases and material presented in non-textual format, as well as large databases such as statistical reports, phone books, governmental statutes, etc. 	<p>Remember: Topical coverage on the Invisible Web is extremely varied and volatile. Some of the information stored on Web-accessible databases may not be</p>

	<p>2. Use to find information that is new and dynamically changing in content, such as:</p> <ul style="list-style-type: none"> o news o job postings o available airline flights, hotel rooms, etc. o stock and bond prices, market averages, etc. <p>3. Using the recommended search engines, type in a keyword placing "database" behind it. EX: In a Google search, type "languages database" or "toxic chemicals database."</p> <p>4. Access search engines to also find subject specific databases: (EX: educational testing database).</p>	<p>substantive or useful. Neither a robot (spider) or human expert is analyzing these pages, so the entire responsibility for evaluation becomes the researcher's burden.</p>
--	--	---

An example of a strategy distinction between search engines and web directories follows:

Begin with a Web Directory if your topic is....	Begin with a Search Engine if your topic is.....
Disabilities	Americans With Disabilities Act
Civil War	Battle of Appomattox
Space exploration	Mars Pathfinder
British literature	Charles Dickens

For an in-depth look at web-based search strategies, visit the following sites:

<http://library.albany.edu/internet/strategies.html>

<http://www.brightplanet.com/deepcontent/tutorials/search/index.asp>

<http://library.albany.edu/internet/started.html>

Evaluation of World Wide Web Resources

You will undoubtedly find many resources pertaining to your topic, however, before utilizing them for research purposes, evaluate the reliability and accuracy of the information. Consider the following methods of evaluation:

1. Author - What are the author's credentials? (think of training, education, background experience). Is the author associated with an institute of higher learning?
2. Sponsor - Who is sponsoring the cost of presenting the information? Is it an educational (.edu), non-profit (.org) or business (.com) organization? Is contact information provided?
3. Visual Layout - Is the format pleasing and text-based rather than full of advertising or outlandish graphics?
4. Purpose - What audience is the author trying to reach? Is the purpose to educate and inform or persuade and entertain?
5. Currency - Is the information up to date? Might additional information supersede the accuracy of the material?
6. Point of View/Objectivity - Is the author presenting both sides of an issue or only using the resource as a means of persuasion? Is the tone academic or emotional and extreme?
7. Support - Does the author provide a bibliography and/or footnotes? Are the citations current?

Learn more about determining the reliability of information at the following websites:

<http://www.lib.duke.edu/libguide/evaluating.htm>

<http://www.library.cornell.edu/olinuris/ref/research/skill26.htm>

World Wide Web Citation Guide

Utilize the following guide as you begin forming citations for web-based (stand alone) resources referenced in research papers or presentations:

APA	National Park Service. (2003, February 11). <i>Abraham Lincoln Birthplace National Historic Site</i> . Retrieved February 13, 2003, from http://www.nps.gov/abli/
Chicago	National Park Service. <i>Abraham Lincoln Birthplace National Historic Site</i> . [updated 11 February 2003; cited 13 February 2003]. Available from http://www.nps.gov/abli/ . <i>Note: Adapted from examples on page 634 of Chicago Manual of Style 14th edition.</i>
MLA	<u>Abraham Lincoln Birthplace National Historic Site</u> . 11 Feb. 2003. National Park Service. 13 Feb. 2003 < http://www.nps.gov/abli/ >.
Turabian	No example in latest edition.

NOTE: Do not use this citation guide for online journals or electronic databases such as NCLIVE. Use this example only for a stand-alone document.

Recommended Sites: A Student Beginner's Guide

CATEGORY	WEBSITES
www Access	http://www.cccti.edu/LRC/Extraneous/Internetaccess.htm
Ready Reference	http://www.refdesk.com
	http://www.ipl.org
News	http://www.wn.com
Student Success	http://www.chemistrycoach.com/linkstoa.htm
Job Search	http://www.roquecc.edu/emp/Resources
Best Of The Web	http://www.nypl.org/links/
Useful Academic Sites	http://www.pasco-hernandocc.com/library/links.htm#arts
E-Mail	http://www.fastmail.fm/
Database For General Academic Research	http://vos.ucsb.edu/
Booklist For Life-Long Learners	http://www.ala.org/yalsa/booklists/obcb
Research Process	http://www.cccti.edu/LRC/infolit/Research%20Process%20Tutorial.htm
Libraries	http://www.cccti.edu/LRC/lrc.htm
	http://www.loc.gov
	http://www.nypl.org
Writing	http://www.unc.edu/depts/wcweb/
Literature	http://www.bartleby.com/cambridge
Literary Index	http://www.galenet.com/servlet/LitIndex/
Literary Criticism	http://www.ipl.org/div/litcrit/
Math	http://www.sosmath.com
	http://www.mathworld.wolfram.com
History	http://www.meredith.edu/history/history-res.htm
	http://campus.northpark.edu/history/WebChron/
Science Search Engine	http://www.scirus.com/srsapp/
Science	http://www.sciencedirect.com
	http://scienceworld.wolfram.com
	http://www.whyfiles.org
Art	http://americanart.si.edu/index3.cfm
	http://witcombe.sbc.edu/ARTHLinks.html
	http://www.biography.com/impressionists/artists_degas.html
Music	http://www.mymusicindex.com
	http://www.music.indiana.edu/music_resources