

# A Review of the World Wide Web for Advanced Legal Research

Prepared by Victoria J. Szymczak,  
Electronic Information Specialist and Adjunct Associate Professor of Law  
for use in Advanced Legal Research, Fall 2001.

---

The topics discussed in this handout can be more fully explored in Jerry Lawson's excellent monograph, "*The Complete Internet Handbook for Lawyers*," published by the American Bar Association in 1999. Loc. 1 Ref., Call No. KF 242 A1 L36 1999. If you are interested in learning more about the history of the Internet, a good, authoritative online source is the home page of the Internet Society at <http://www.isoc.org>.

## BASIC INTERNET TERMINOLOGY

There are many "technical terms" you will encounter in the cyber world. You should become familiar with some dictionaries to help you when you do not understand a concept. If there are any technical terms in this handout with which you are unfamiliar, look them up! Listed below are some online glossaries which you might want to consult:

Sharpened at <http://www.sharpened.net/glossary/index.php>  
U Geek Glossary at [http://www.ugeek.com/glossary/glossary\\_search.htm](http://www.ugeek.com/glossary/glossary_search.htm)  
Web Guest at <http://www.webguest.com/glossr.html>  
What Is? at <http://whatis.techtarget.com>

## ***E-mail***

E-mail is the most common use of the Internet, even though it is not the most glamorous. It is important for business and personal communication. It allows messages or files to be sent to other people. Court documents, for example, are being filed through electronic mail with greater frequency. Attorneys share documents through secured E-mail transmission as a matter of course. Law professors communicate with students using this medium. The standard format for E-mail addresses is name@site. For example,

[vszymc@brooklaw.edu](mailto:vszymc@brooklaw.edu)

Any E-mail message that you receive or send will have two distinct parts: the header and the body. The header of an e-mail message contains useful information, such as the name of the sender of the message, the name of the recipient, the subject of the message, whether any carbon copies were sent (and to whom), and the time at which it was sent. The body of an e-mail consists of the actual contents of the message.

*This is an example of an E-mail header:*

O-SMTP-Date-posted: 6-Oct-99 10:36:49 -0500  
From: APIG @ SMTP2 (Arnold the Pig) {apig@greenacres.com}  
To: VSZYMC @ BROOKLAW  
Subject: DINNER TONIGHT  
Date: 6-Oct-99 9:34:31 -0500  
O-SMTP-Envelope-From: <apig@apig@greenacres.com>  
Via-host: BROOKLAW.BROOKLAW

The transfer of information via E-mail often involves the use of attached files to the body of the message. An attached file is commonly referred to as an "attachment." An attachment can be in many formats such as audio, text, video, graphic, or even a virus. In order to extract an attached file it is important that you have the correct application to open it. For example, if the attachment is a Word Perfect document, the recipient must have a word processing program on his or her computer that can "read" Word Perfect in order to view the attachment. Usually a recipient can determine the file format by looking at the three letter extension in the file name that follows the period. See File Formats for more information.

If your operating system (i.e., Windows 98) does not recognize the file format, it will ask you which program you would like to use in order to view the document. This is a good indication that you do not have the appropriate program installed on your computer. There are two possibilities you might consider if you find yourself in this situation:

- Some programs allow you to download a "viewer" to view documents created with their program. The viewers are usually free. See File Formats for more information.
- You can download an image viewer that will execute when you receive a incompatible file format. Check to see if the image viewer will support the types of files you need to open. The three suggestions below are general purpose viewers that should support more file formats:
  - Polyviewer at <http://www.polybytes.com> > Downloads
  - Dave Central - SBJV Image Viewer at <http://www.davecentral.com/2289.html> (in addition to this viewer, "Dave" has several other more specific viewers on his web site.

If a file is larger than 2 MB, than it is inappropriate to send it as an E-Mail attachment. It could clog up your IP Host's E-Mail server, and slow down your network.

Large files and programs are transferred using the File Transfer Protocol service on the Internet. This requires additional software and a rudimentary understanding of command protocols. If you find yourself sharing large files with others on a frequent basis, you may be best served by registering for a file hosting service. There are many good, free hosting services on the Internet. A good list of free hosting sites can be found at <http://webwizards.net/useful/wbfs.htm>. A couple of suggestions are:

- Free Drive at <http://www.freedrive.com>
- Yahoo Briefcase at <http://briefcase.yahoo.com>

### ***Electronic Discussion Groups***

An electronic discussion group is a mailing list for discussions on a particular topic. It can also be referred to as a mailing list, e-group, forum, e-conference or a listserv. You use e-discussions to stay current in a particular area and/or be part of an electronic support group to whom you can turn to ask for help. ***This can be an important research tool if you find one relevant to your paper topic.***

An e-discussion is run by a software program. LISTSERV is one type of electronic discussion group software (this particular program is so popular and has been around so often that people use the word listserv to describe any electronic mailing list). Other significant software programs include LISTPROC and MAJORDOMO. The type of software you use, will determine which commands you will use to subscribe, unsubscribe, post messages, answer messages, and execute other commands. Some commands that are common to mailing lists are to request a list of subscribers, receive a digest of list messages, and postpone deliveries. Regardless of the type of program being used, all E-mail discussion groups have two addresses associated with them:

- A "listname address: which is the address to which you send any messages that you would like read by other list subscribers, and
- An "administrative address," which is the address to which you send any commands or requests that affect your subscription to this list.

There are several ways to find an e-discussion on a variety of topics in the practice of law. You might be advised by word of mouth. You can also consult any number of print and online directories. On the Internet, you will want to run searches on the following sites to find a group of interest to you:

Lyonette Louis-Jacques, *Law Lists*, at <http://www.lib.uchicago.edu/~llou/lawlists/info.html>, and

In addition to law related e-discussions, you might want to join a non-law related listserv. There are lists of general discussion groups too. You can consult the following collections:

Listz at <http://liszt.com>

Tilenet at <http://www.tile.net/lists>

### **World Wide Web**

Although E-mail is the most utilized service on the Internet, it is the World Wide Web (WWW) that is the most recognized. The advantage of the WWW is that documents are linked to each other in a comfortable Windows environment. The type of material you can access is “virtually” limitless -- as long as you have a capable computer, an Internet Service Provider, and a browser.

Material available to you on the Internet ranges from a simple text file (i.e., a case from the Supreme Court) to audio files (i.e., the oral arguments heard at the Supreme Court) to short movies broadcasted via the Internet (sorry, the Supreme Court justices are not yet ready for Hollywood). There are also subscription databases that provide better search capabilities and high quality editorial enhancements for a fee. Lexis and Westlaw are typical examples of subscription databases available on the web.

To access the WWW, you use a “browser.” The three major browsers in the United States are Netscape, Internet Explorer (IE) and America Online (AOL), though these represent only three of several dozen browsers (see Browsers Archive at Evolt, <http://browsers.evolt.org/>. ) America Online is a special browser. It channels information to you via its home page and does not function the same way that Internet Explorer and Netscape function. Experienced web surfers tend to shy away from AOL because it is a little more restrictive than a true browser like Netscape and Internet Explorer.

Generally, for the end-user, IE and Netscape function similarly. There are some differences in terminology, but the main difference is simply in the display of information on the web. A web page might look different in IE than it does in Netscape. You can test this by looking at the Brooklyn Law School web page in both IE and Netscape. Often, you will see a notice on a web page that it was written for, or is best viewed using, one or the other browser. For this class, we will be looking at the web through the eyes of Netscape Communicator, version 4.7. Publishers produce manuals for each version of Netscape. For a more detailed account of version 7 and Netscape generally, you can consult :

*Netscape Communicator 4.7: Quick Source Reference Guide*, Quick Source (2000)

Patrick Clancey and Rebecca Nelson, *Netscape Communicator Essentials*, Que (1999)

Rob Tidrow and Greg Robertson, *The Essential Netscape Communicator Book*, Prima (1997)

## ADDRESSES AND DOMAIN NAMES

PART 2

PART 4

<http://ci.nyc.ny.us/html/ccrb/html/file.html>

PART 1

PART 3

A URL has a definite structure that consists of 4 parts: transfer protocol, servername.domain, directories subdirectories and filename type. Familiarity with URL structure will help you evaluate your cyber resources, which is a very important aspect of computer assisted legal research --- *and your final research guide!*

URLs are not made up. You must apply for one when you start up your web site -- unless you are paying another company to host your web site. Applications for a domain name are made to a select group of networking companies around the world. In the United States, the Internet Corporation for Assigned Names and Numbers (ICANN at <http://www.icann.org>) is the non-profit corporation that is responsible for domain name system management and IP address space allocation. This organization accredits companies that seek to serve as a U.S. domain name registry. Once you apply for a URL and pay your fee to an accredited company, a twelve digit numerical web address (IP address) is assigned to you. A verbal equivalent will be adopted as an alias. This is the URL with which most people are familiar. The basic necessary format is for a URL is:

transfer protocol://servername.domain

### ***Part 1 - Transfer Protocol***

The transfer protocol, or the first portion of a URL, indicates which type of Internet service you will be using. For example, <http://> indicates you will be accessing a world wide web site. Other services you may encounter are gopher:// to indicate a gopher site, <ftp://> to indicate an ftp service, <mailto://> for E-mail, and telnet:// for telnet. These notes refer to the world wide web and URLs that begin with <http://>.

### ***Part 2 - servername.domain***

The string of words following <http://> are referred to as the server name and top level domain name. They are typically divided into three groups. The first two parts of

the address are referred to as the server name. Often, but not always, a company will call its server for the world wide web “www” and follow it with the name of the institution.

For example: [www.brooklaw.edu](http://www.brooklaw.edu)

The second part of a URL is very important for researchers to understand. It is the key to understand from where the information is coming. It will also assist you in your evaluation of the information presented on that site.

In the early 1990's, there were only six types of domain names, and the majority of web sites were hosted in the United States. These six names are referred to as *generic top level domain names (gTLDs)*. They include: .edu, .gov, .com, .mil, .net, and .org. This portion of the URL is meant to tell you what type of company or institution you were contacting. They are shorthand for the following types of institutions:

- .edu = educational
- .gov = government
- .com = commercial
- .mil = military
- .net = networking organization
- .org = nonprofit organization

The Internet is more crowded and complicated than when these six domain names were first put into circulation. Today, there are many other types of entities that do not fit neatly into one of these six categories, and more than 7 million of web sites hosted all over the world. Later this year, additional top-level domain suffixes will be approved by the organization(s) that coordinate the assignment of identifiers to individuals and groups around the world. These TLDs are:

- .aero = air transport Industry
- .biz = businesses
- .coop = non-profit cooperatives
- .info= unrestricted use
- .museum= museums
- .name = individuals
- .pro = accountants, lawyers, and physicians

\*.int = international organization created by an international agreement (This is not an “approved” top-level domain, but it is used frequently enough that you may see in when conducting your research.)

Other initiatives already in place include geographic designations in domain names (ccTLDs). The structure of the URL using geographic designations incorporates country, state, county and city indicators. You will continue to see the initial six generic names, plus the seven new TLDs, but you will increasingly come across a two letter code at the end of the URL that will tell you from where the site originates (or is hosted). For example, .us indicates the site is hosted in the United States while .ca indicates the site is maintained in Canada. For a complete listing of country codes see IANA CCTLD at <http://www.iana.org/cctld/cctld-whois.htm>.

Within each country, there will be a further breakdown according to state, territory, county, or city. So, the abbreviation .ca can also indicate a Californian web site when it is followed by .us For example, <http://fire.state.ca.us> indicates a Californian web site while <http://fire.ca> indicates a Canadian web site. The geographic domain naming structure also includes allowances for special types of organizations such as schools educating students in the K-12 levels, special federal agencies and other designations as outlined on the US Register Site - RFC 1480 at <http://www.us>.

Currently, in the United States, the federal and state governments are the only serious users of ccTLDs. The standard format for state governments is:

[part1].state.[postal abbreviation].us

or

<http://www.state.ny.us> (using NYS as an example.)

To obtain more information about domain name structure, including the legal issues attendant to this process, you should consult:

- a) Internet Assigned Numbers Authority (IANA) at <http://www.iana.org>
- b) Internet Corporation for Assigned Names and Numbers at <http://www.icann.org>
- c) InterNIC Website at <http://www.internic.net>
- d) Official Registry of US Top Level Domains at <http://www.us>
- e) World Intellectual Property Organization Internet Domain Process at <http://wipo.int/index.html.en>
- f) The United States National Telecommunication and Information Administration at <http://www.ntia.doc.gov>

- g) CORE (Internet Council of Registrars at <http://www.corenic.org>)

### ***Part 3 - Directories and Subdirectories***

The third part of a URL indicates the directories and subdirectories on the server. Just like your word processor, a server stores web pages in directories to keep them organized. The path of directories and subdirectories are separated by a forward slash (/). Some web sites do not allow you to go directly to a web page in one of their subdirectories. If you try to go to it, you will be redirected to the entity's home page or denied permission to enter completely.

### ***Part 4 - File Types***

The last part of the URL specifies the individual document at which you are looking. In addition to the name of the document, you can also tell what file type it is. The standard file type for the World Wide Web is .htm or .html which stands for hypertext markup language. Documents ending in .htm or .html do not require anything more than your standard browser to view it. Other types of files will require the use of a helper application or plug in. When you encounter such a file, your browser will provide further information.

A popular type of file used in the legal environment that requires the use of a helper application is .pdf. This acronym stands for portable document format. A PDF document requires you to have an additional application on your hard drive called Adobe Acrobat Reader in order to view it. This application can be downloaded for free from the Adobe web site at <http://www.adobe.com>. The advantage of this format is that it allows an image perfect representation of a document to appear on your screen and print on your printer. This is the reason it is so popular with lawyers. It permits you to provide accurate citations in your briefs and other legal documents. It is also more difficult for miscreants to tamper with PDF files than simple HTML documents. This makes them inherently more reliable. A significant disadvantage of using PDF files on the Internet exists for Internet researchers. They cannot be indexed by search engines, making them inherently more difficult to find than other types of file formats. See Search Engines, Invisible Web for more information about this frustrating aspect of Internet research.

Other common file types who may run across on the Internet are:

- a) Formats for static visual images – .gif, .jpg, .bmp
- b) Formats for moving images - .mov, .avi
- c) Formats for audio files – .ra, .wav, .mpg

- d) Compressed file formats (you will need special software to decompress these types of files if you want to download them to your computer) – .zip, .tar
- e) Viruses - .vbs

## **GETTING STARTED WITH SEARCH ENGINES, SUBJECT DIRECTORIES & SOME GOOD WEB SITES**

Unless you already know where on the Internet you can find the information you seek, you will have to do some research to obtain the URL of a web site that houses that information. It is very common for people to go to a popular search engine or directory and insert one or two words into a search box, cross their fingers, and hope they find something useful within the first hour of surfing! This may work for you, especially since search engines and subject directories are getting better every day. However, there are some preliminary strategies that you can use that might be more useful.

### ***GUESS!***

Although this might not sound like the most intelligent way to approach the Internet, it can be very effective. Identify the information you need and think about what organization would be most likely to post that information on their web site...and then guess their URL! The URL will most likely begin with www. It will then be followed by the name or acronym of the institution you want to contact. The URL will end in either a gTLD or will take the form of the geographically oriented domain naming structure.

### ***Print Directories***

If you are not guessing correctly, it is a wise idea to consult resources that present a more definite structure. One such tool is a current, hard copy directory of web sites. Some suggested titles in the Brooklyn Law School Library include:

Joshua D. Blackman, *The Internet Fact Finder for Lawyers*, Chicago, Ill. : American Bar Association, Law Practice Management Section, 1998. Loc. 1 Ref. Call No. KF 242 A1 B55 1998.

Diana Botluk, *The Legal List*, St. Paul, MN : West Group, 1999. Loc. 1 Ref., Call No. KF 242 .A1 L44 1999.

Yvonne J. Chandler, *Guide to Finding Legal and Regulatory Information on the Internet*, New York, NY : Neil - Schumann, Loc. 1 Ref., Call No. KF 242 A1 C48 1998.

T.R. Halvorson, *Law of the Super Searchers : The Online Secrets of Top Legal Researchers*, Medford, N.J. : CyberAge Books, 2000. Loc. Main, Call No. KF242.A1 H35 2000

Erik J. Heels and Richard P. Klau, *LAW LAW LAW on the Internet*, Chicago, Ill. : American Bar Association, Law Practice Management Section, 1998. Loc. 1 Ref., Call No. KF 242 .A1 H33 1998.

Judy A. Long, *Legal Research Using the Internet*, St. Paul, MN : West Group. LOC 1 Ref., Call No. KF 242 .A1 L66 2000.

Don MacLeod, *The Internet Guide for the Legal Researcher*, Teaneck, N.J. : Infosources Pub, 1996 (Supp. Winter 2000), Loc. 1 Ref., Call No. KF 242 .A1 M34 1997

Antje Mayes, *Legal Research on the Internet : a Compendium of Websites to Access United States Federal, State, Local, and International Laws*, Buffalo, N.Y. : W.S. Hein, 1999, Loc. Main, Call No. KF242.A1 M39 1999

Herbert N. Ramey and Samantha A. Moppett, *Navigating the Internet : Legal Research on the World Wide Web*, Littleton, Colo. : Fred B. Rothman & Co., 2000, Loc., Main Call No. KF242.A1 R35 2000

### **Library Catalogs**

Libraries, including Brooklyn Law School Library, are cataloging web sites. This is important for the researcher for two very important reasons. First, a catalog permits you to make use of a structured classification system (in direct contrast to the lack of structure on the Net). Second, a resource is not included in an academic catalog unless it has been thoroughly reviewed for content, reliability, and stability.

A web site will be included in a results list generated by you query. If you are using the web version of Brooklaw, you will be able to link to that site directly from your catalog results. So, for example, if you conducted a word search for "sustainable development" on our catalog, your first hit might be a web site, followed by a report, a monograph, an U.N. publication, and so on. If you wanted to retrieve only web sites, you can conduct a subject search for "internet file," and then select LIMIT to restrict your results to a particular topic.

A list of United States law school library catalogs and their URL's can be found at the Washlaw web site (<http://www.washlaw.edu/lawcat.lawcat.html>). You should

consult catalogs of large research libraries or schools that have specialty collections. For example, Franklin Pierce law school hosts a critically acclaimed environmental law program. The library collection that supports this program is rich and diverse. You would service yourself by browsing its collection if you are researching an environmental topic.

### ***Search Engines v. Portals, Megasites or Subject Directories***

There are approximately over a billion of indexable individual web sites which comprise the Internet. Eventually, there will come a time where you will need to consult a search engine or subject directory on the Net to sort through those web sites. The main difference between the two is that a search engines are organized according to a computer algorithm, while subject directories are organized and reviewed by people. A portal or megasite are simply large subject directories devoted to specific subject areas.

Both Search engines and subject directories rely on the use of a computer program, called a spider or robot, to “crawl” over the Internet, identifying new web pages, and storing those web pages in a database. The database is what you are searching...not the Internet as it exists in the moment! Search engine databases often provide basic indexing, and are usually larger than a subject directory. As of April, 2001, the largest general purpose search engines were WiseNut, Google, and Fast.

Subject directories are smaller than search engine databases for a few reasons. First, directories often index only the top level of a web site. This is exactly the opposite purpose of a search engine which indexes every link on a page. Second, classification of web pages at a subject directory service involves divine human intervention. The efforts of good web site classifiers make your search results at subject directories much more accurate.

The differences between search engines and subject directories will lead you to use one over the other depending on the type of information you seek. Typically, a *search engine* should be used when you are trying to find a specific document. *Subject directories* are better suited to obtain an overview of a topic. Many “*specialty*” sites are accessible on the Web. Specialty sites focus on a particular topic (i.e., the law). Modern versions of search engines and subject directories will offer the opportunity to access the web both ways. Regardless of which type of search device you use, the hallmarks of better research tools include a relevancy ranking or sorting capability, ability to conduct advanced searching, and annotated web sites.

### ***Megasites or Portals for Legal Materials***

A good starting point to find web sites pertaining to your research, is to make use of a megasite or portal. A portal is an extensive collection of links dedicated to a particular issue or area of concern. It is meant to be a one-stop shop for you to access any and all information in a given field of interest. The scholarly pursuit and practice of the law is well represented by many such sites. There are several law school

sponsored portals that are rich in content and well maintained. It is a good idea to consult them when conducting legal research on the web. Some suggested web sites include:

1. School sponsored sites:
  - a. Cornell University's Legal Information Institute at <http://www.law.cornell.edu>
  - b. JURIST: The Law Professor's Network -- Subject Guides at [http://www.law.pitt.edu/subj\\_gd.htm](http://www.law.pitt.edu/subj_gd.htm)
  - c. Washlaw Web of the Washburn School of Law at <http://www.washlaw.edu>
2. Commercial sites:
  - a. Catalaw at <http://www.Catalaw.com>
  - b. Findlaw (West's Free Site) at <http://www.findlaw.com>
  - c. Hieros Gamos at <http://www.hg.org>
  - d. LexisOne (Lexis's Free Site) at <http://www.lexisone.com>

### ***Law Related Research Guides***

In addition to these wonderful megasites, educators and practitioners have developed research guides that help researchers through the process of researching specific topics within the law. These guides are also rich in web links.

- a) LLRX at <http://www.llrx.com/>
- b) Legal Information Institute Law About ... at <http://wwwsecure.law.cornell.edu/topics/topic1.html>
- c) LexNotes at <http://www.lexnotes.com>
- d) New York University Guide to Foreign and International Law Sites at [http://www.law.nyu.edu/library/foreign\\_intl/index.html](http://www.law.nyu.edu/library/foreign_intl/index.html)
- e) Jurist Law at <http://www.jurist.law.pitt.edu>

### ***Specialized Search Engines***

There are hundreds of general search engines on the Internet. Within the last couple of years, search engines and subject directories devoted to a particular topic have sprung up. Typically, these speciality sites, or portals, use the search engine software developed by a general search engine. They are subject to the same limitation and advantages as that of general search engine sites such as Google.com or WiseNut.com.

- a) Law Runner at <http://www.lawrunner.com>
- b) Law Crawler at <http://lawcrawler.findlaw.com>
- c) Law Researcher at <http://www.lawresearch.com/vs/search10.htm>

### **Megasearch engines**

A megasearch engine acts as a gateway to other search engines. You insert your query and it will search a number of other search engines and, usually, sort them by relevance under each database queried. There is one that does this for law oriented search engines. It is called Law Guru located at <http://www.lawguru.com/search/research.html>. Law Guru claims to search 500+ legal search engines and tools. Be cautious when using a megasearch engine of any type. Typically, your advanced searching techniques will be stripped down by the megasearch engine so that only the most simple of search techniques will be recognized. Since you are submitting your query to multiple search engines, all with their own special search language, advanced commands will not be recognized. Keep this in mind if you choose to use this type of resource.

For a discussion of general web searching techniques, refer to pp. 1 to 31 in Randolph Hock, *The Extreme Searcher's Guide to Web Search Engines*, Medford, NJ : CyberAge Books, 1999. A more extensive review of web searching can be found in T.R. Halverson, *Law of the Super Searchers : The Online Secrets of Top Legal Researchers*, Reva Basch, ed., Medford, NJ : CyberAge Books, 1999. Two excellent online references for more detailed information about search engines are: Danny Sullivan's Search Engine Watch at <http://www.searchenginewatch.com> and Greg Notess's Search Engine Showdown at <http://www.notess.com/search>

### **"INVISIBLE WEB" Engines**

Like space, the Internet has its own black hole known as the "invisible web." The term refers to that material that a search engine cannot index, thus the term "invisible." There are many reasons, known and unknown, why general-purpose search engines cannot see these black holes. A fundamental problem with search engines is that they index pages that other pages point to. So, if there are no links pointing to a page that a search engine spider can follow, it cannot be indexed. It becomes part of the invisible web. Other reasons for the existence of the invisible web include the type of data stored at a web site. A search engine cannot index graphics, CGI scripts, Macromedia flash or, except for Google.com, PDF Adobe files. Search engines cannot index material stored in databases. For example, you cannot search Google.com for materials held in the Brooklyn Law School Library. The database that keeps track of our holdings cannot be read by a search engine.

It is estimated that 74% of the Internet falls into the invisible web. Some examples of the types material not crawled by general-purpose search engines include the databases that hold IRS Forms, FBI Crime Statistics, SEC corporate securities

filings, and the patent database at the U.S. Patent Office and IBM. In fact, the entire GPO Access web site ([http://www.access.gpo.gov/su\\_docs/aces/aaces002.html](http://www.access.gpo.gov/su_docs/aces/aaces002.html)) and its related web sites are not indexed by search engines. This includes the United States Congress and Library of Congress. For a legal researcher, it is important to note that many primary law, statistical, and periodical databases cannot be seen by general or specific purpose search engines. A very small collection of special search engines that can puncture the law-related invisible web sites is represented below.

- a) Gary Price's Direct Search at <http://gwis2.circ.gwu.edu/~gprice/direct.htm> (Mr. Price maintains a subject oriented approach to the invisible web. His collection on legal materials is found at <http://gwis2.circ.gwu.edu/~gprice/direct.htm#Legal> )
- b) The Invisible Web at <http://invisibleweb.com/> (also a subject oriented web site)
- c) AlphaSearch at <http://www.calvin.edu/library/searreso/internet/as> (pull down subject menu has an entry for Law)

### **SEARCH TIPS** (See *Quick Reference Chart at the end of this handout.*)

There are hundreds of search engines on the Web, but there are millions of web pages. A serious research session on the Internet requires you to query several content rich databases to obtain as many relevant hits as possible, since the databases at the web sites are not identical. All search engines are unique and have different searching parameters. The search logic is not standardized on the Internet. For effective powerful searching you need to understand the logic used by search engines. The top search engines provide good help screens and sample searches to get you started. Despite the different search methods, there are some generalities with which you should become familiar. These are noted below.

### IDENTIFY KEYWORDS

When conducting a search, break down the topic into key concepts. If any one concept is more important to you than others consider placing it first in your search string. For example, to find information on the impeachment proceedings against Richard Nixon following the Watergate scandal consider isolating the following terms.

### NIXON WATERGATE IMPEACHMENT

### TRUNCATION

Truncat\* - \* helps you find variations of root words. Use an asterisk symbol (\*) to retrieve various word forms. The asterisk (\*) symbol tells the search engine to return alternate spellings for a word at the point that the asterisk appears.

Impeach\* will yield impeachment, impeachable, impeached, impeached.

## PHRASES

Phrase searching often requires you to surround a group of words with double quotes. It tells the search engine to only retrieve documents in which those words appear side-by-side. Phrase searching is a powerful search technique for significantly narrowing your search results, and it should be used as often as possible.

"richard milhous nixon"

"impeachable offense"

## CAPITAL LETTERS

Most search engines interpret lower case letters as either upper or lower case. It is good practice to always put your search terms in lower case. If you purposely want to limit your results to capital letters (e.g., Richard Nixon) you can enter it that way for those results.

### ***Boolean Searching***

Boolean searching uses connectors such as AND, OR and AND NOT. The use of connectors can make your searching more precise than simply entering in a set of terms. In addition, some search engines will insert an automatic AND between any words that appear in the query box. It should be noted that professional searchers find search math type queries are equally effective as boolean searching.

Using AND helps you narrow your results. Connect your terms with AND if you want the search engine to retrieve web pages containing all the words. The search engine will not return pages with just the word Nixon. Neither will it return pages with the words Nixon and Watergate. All three words must be present.

nixon AND watergate AND impeachment

Using OR helps you expand your results. When OR is used, the search engine may return pages with a single keyword, several keywords, or all keywords. To get the best results, it is good practice to put your terms connected by the OR in parenthesis.

(Nixon OR president) AND (scandal OR Watergate) AND impeachment

Using AND NOT helps you eliminate false hits. If your initial result includes too many unwanted results, try using the and not command to narrow your results.

impeachment AND NOT clinton

### ***Power / Advanced Searching***

Power searching allows you to manipulate your search query in very specific ways. This type of searching is akin to field searching in Westlaw, or segment searching in Lexis. Depending on the search engine you use, this type of refinement may be offered on a pull-down menu, or taken advantage of only by command. The most relevant power search command for legal researchers is the SITE SEARCH.

A particularly useful advanced search feature permits limiting your search to a specific domain. This is advantageous when a web site does not have a search feature, or their search engine is does not work well. For example, to find specific documents at the FBI for documentation on their E-mail sniffer called CARNIVORE you can use the following command:

**url:fbi.gov and carnivore**

The above search can be used with the following search engines:

- Alta Vista at <http://www.av.com>
- Fast (Advanced Search Only) at <http://www.alltheweb.com>
- Lycos Advanced Search at <http://www.lycos.com>
- Norhtern Light at <http://nlsearch.com>

In Google.com, you must use the search: site:fbi.gov and carnivore!

Other powerful search features to keep in mind include

- The ability to SORT your results.
- The opportunity to retrieve results in a specific language.
- A date restricter that limits results by the date a document was loaded onto a web site (NOT the date of the document).

### ***Tutorials***

There are many online tutorials that will take you from the basics of web searching to more advanced techniques of web searching.

- Web Searching, Sleuthing and Sifting at <http://www.thelearningsite.net/cyberlibrarian/searching/ismain.html>
- Mining Company's Web Search Guide <http://websearch.about.com/cs/howtosearch/index.htm>
- Finding Information on the Internet : A Tutorial from the Teaching Library Workshop at the University of California a Berkeley <http://www.lib.berkeley.edu/TeachingLib/Guides/Internet/?findInfo.html>

## **CRITICAL THINKING : EVALUATING A WEB PAGE**

Researchers should evaluate the quality of a web page if they are going to rely on the information that they retrieve off of the Web. If we were conducting our research using print sources, we would automatically ask ourselves several questions. The first question legal researchers generally ask is "is it current?" Often this is answered by looking on the front of a pocket part, or finding the publication date of the latest supplement provided by the publisher. Those familiar with print material might also ask "who is the publisher?" This is because they are more familiar with the names involved in legal publishing and realize that some are better arranged, some are better organized, and some serve as more persuasive authority than others.

There are several factors that play into your evaluation of a web page. Many of them do not differ from your consideration when evaluating and accessing a hard copy resource. Others are unique to the World Wide Web. In order of discussion these factors are authority, currency, objectivity, depth and ease of use. *This information should be included in annotations for web sites in your research guide.*

### **Authority**

The author of the page should be a known, stable organization, or individual, whose work can be relied upon. This information should be readily available for someone to review at the web site. Government organizations, schools and universities, and even some commercial entities maintain good, reliable sites that can be trusted. Many law schools maintain information and links to a large variety of legal resources. Two good examples of this type of endeavor are WashLaw from Washburn University at <http://www.washlaw.edu>, and Cornell University Legal Information Institute at <http://www.law.cornell.edu>.

Keep in mind the domain name structure when trying to authenticate information on a web site. The type of domain name can provide clues about who is providing the information. The top level domain name for Cornell and Washburn both end in .edu. Information from educational institutions is quite reliable. This is also true for material at an official United States government web site designated with the suffix .gov. It becomes more difficult when the top level domain ends in .com or .org because these designations are available to the public-at-large. They are NOT tied to any specific type of entity. Consider these three possible web addresses for the White House:

<http://www.whitehouse.com>

<http://www.whitehouse.net>

<http://www.whitehouse.gov>

Evaluation of the URL become particular important when you are looking at a pure text document without any identifying clues concerning its authenticity. It is not, however, the most reliable tool you can use to ascertain the reliability of a web site. For example, a .com web site may be completely unbiased and a terrific source of information. Frequently, experts in a field of law may maintain their own web site which uses a .com ending. For example:

<http://www.mexlaw.com/> ,  
sponsored by Professor Jorge Vargas

The Mexlaw web site is sponsored by a Professor. Included on this web page is information about him and how to contact him via E-Mail. To confirm the information Professor Vargas provides on his web page, you should conduct an Internet search using his name and virtually visit the organizations to which he is affiliated. With the new domain suffixes, Professor Vargas may opt to use .prof or .name instead of, or in addition to, .com. Examination of a URL should always be followed with a review of who is providing the information.

### ***Currency***

If you are searching for the most current information, you will want to have a very current site. Good sites will often post the date that the content was last updates. Alternatively, a significant amount of information, including the date the document was last updated, can be obtained by using the pull-down menu called OPTIONS. If you click on this feature and then select Document Info, you will be shown a screen that provides the viewer with information about the document. It includes a last updated field. To return to the web site that you were viewing, close the window by clicking on the X in the box located in the upper right hand corner.

Remember, though, currency may not always be important in the case of archival or static information, such as a government report.

### ***Objectivity***

Consider who the author is when reviewing the information. Depending on their agenda, you may want to reconsider the reliability of the information in light of their bias. For example, a page might be sponsored by a marketing company whose objective is not to provide you with good, reliable information, but to sell you a product. You should think twice about relying on the product safety statistics posted on this type of web site. In the legal environment, you should be especially cautious about relying on information produced by special interest sections that might offer slanted information.

When pursuing a topic for legal research, a web site maintained by a law firm or lawyer specializing in the field may provide some basic groundwork for the novice. Additionally, organizations rallying behind a particular issue can also be an important

source of information although you should be careful if they present a biased view of the topic. You must always review the information being presented. For example, consider these two web sites concerning the General Agreement on Tarriff and Trade and the World Trade Organization:

<http://www.gatt.org>

<http://www.wto.org>

### ***Depth***

Many pages are just one-stop shops for the legal researcher. These types of mega-pages do not provide content, but instead provide links out to other content-based web sites. In this type of situation, the web presence serves more as an indexer for other sites, than an actual information provider.

### ***Ease of Use***

If you are confused about how to maneuver within the web site you should be cautious about recommending it to other researchers. It is frustrating to go through several pages of information trying to find your way back to your entry point. A well-designed web site will provide links back to their main home page on all of their sub-directory pages that you view. If you find yourself in a poorly designed web site, remember to use your GO feature to get yourself back to a solid point of reference.

Another concern for the surfer revolves around special software requirements needed in order to view the information. Often, these software requirements take the form of "plug-ins."

## **HOW TO CITE RESOURCES FOUND ON THE INTERNET**

### ***Blue Book Format***

The controlling rule in *The Bluebook Uniform System of Citation*, 18<sup>th</sup> ed. for citing to digital material is Rule 18. Rule 18.1 addresses commercial databases such as Lexis, Westlaw, and First Search. Rule 18.2 focuses on the World Wide Web exclusively and does not generally include citation formats for the other services of the Internet such as E-Mail, FTP, or Usenet. The *Bluebook* states that:

“The principles of Internet citation ...are intended to be flexible guidelines applied in light of the nature of the information being cited, the state of the Internet as it develops over time, the standards or requirements applicable to the citing work, and the uses to which the citing work are to be put.”

Regardless of the acceptance of the Internet as a resource needing a uniform citation format, Rule 18.2.1 still requires citation to printed sources “except when the information is not available in a printed source, or if the traditional source is obscure or hard to find AND when the citation to an Internet source will substantially improve access to the same information in the traditional source.” (Emphasis added.) The rule presents the following three scenarios.

A. When we are using the Internet cite as a parallel citation, and have consulted the corresponding print resource, we must use the explanatory phrase “available at.”

Example: *Bush v. Gore*, 531 U.S. \_\_\_\_ (2000), *available at* <http://laws.findlaw.com/us/000/00-949.html>

B. For material available exclusively on the Internet, use the explanatory phrase “at.”

Example: J.T. Westermeier, *Ethical Issues for Lawyers on the Internet and World Wide Web*, 6 RICH. J.L. & TECH. 5, ¶7 (1999), at <http://www.richmond.edu/jolt/v6i1/westermeier.html>.

C. When we are using the Internet cite in place of a print resource cite, and have NOT consulted the print resource, we do not use any explanatory phrases.

Example: *Bush v. Gore*, 531 U.S. \_\_\_\_ (2000), <http://laws.findlaw.com/us/000/00-949.html>

### ***Vendor Neutral Citation Format***

The American Bar Association and the American Association of Law Libraries have also addressed standards for Internet citation. These organizations are working to create a “vendor neutral” citation format that will be applicable for both print and Internet material. Currently, the states that are permitting use of these standards are Arizona, Colorado, Louisiana, Maine, Mississippi, Montana, New Mexico, North Dakota, Oklahoma, South Dakota, and Wisconsin. The use of a public domain format citation for states that provide vendor neutral citation is also covered in Rule 10.3.3 of the Bluebook.

*Judicial decisions* under the Uniform Citation Format take the following structure:  
[year of decision] [postal abbreviation] [chronological number of the case]  
[paragraph number for pinpoint cites]

To illustrate: *Reds v. Mets*, 1999 US 163

*Statutory material* under the Uniform Citation Format take the following structure:  
[name of statute][standardized code designation if available] [numbering of the code] [the current through date]

To illustrate: MS Code Sec. 81-9-107 (1998 through Reg Sess)

To obtain more information about the Uniform Citation Standards visit the American Bar Association Web Site at <http://www.abanet.org/citation> and the American Association of Law Libraries Web Site at <http://www.aallnet.org> or consult The Universal Citation Guide, Committee on Citation Formats, American Association of Law Libraries, Madison, WI : State Bar of Wisconsin, 1999, LOC. Res., Call No. KF245 .U58 1999.

### ***Other Online Citation Guides***

Beyond standards required in legal citation format, there are many style guides for other types of digital material. These might be consulted if the standard format for legal sources does not meet your needs. For example, the Blue Book does not cover citation to E-Mail communication; however, other citation formats, such as the American Psychology Association (the most popular citation format in American colleges and universities), do provide guidance for this service.

- Walker, Janice R, *The Columbia Guide to Online Style*, New York : Columbia University Press (1998). Also, see their accompanying web page at [http://www.columbia.edu/cu/cup/cgos/idx\\_basic.html](http://www.columbia.edu/cu/cup/cgos/idx_basic.html)
- APA Electronic Reference Formats at <http://www.apastyle.org/eleceref.html>
- The MLA Style to Online Citation at [http://www.mla.org/main\\_stl.htm](http://www.mla.org/main_stl.htm)

**QUICK REFERENCE CHART  
FOR BOOLEAN SEARCHING**

	A space between words means AND	A space between words means OR	Phrase searching can be done in:	Mine a site using this command
Alta Vista		X	X	url:[add url]
Fast	X		X	url:[add url] (Adv. Search)
Excite		X		
Google	X	X (must be in upper case)	X (Adv. Search)	site:[add url]
HotBot	X		X	domain:[add url]
Lycos	X		X (Adv. Search)	url:[add url] (Adv. Search)
MSN	X		X	domain:[add url]
Northern Light	X		X	url:[add url]
Teoma	X		X	
WiseNut	X			