INTERNET BASICS AND EVALUATION

WHAT IS IT?

The Internet is a system through which information flows – it is NOT the information itself.

Internet Explorer and Firefox are “browsers” – programs (in your computer) that enable computers to “talk” to each other, and let you use your computer to view the Web.

The Web (World Wide Web: www.) is HUGE – there are BILLIONS of documents available. The Web is a means of linking to many individual pages that people, companies, or organizations have put on servers. The amount and variety of material available on the Web range from absolutely useless, at times even dangerous information, to newsworthy, helpful, and important material.

You DO NOT search the Web directly, but through various searching databases -- called search engines or search directories

Some definitions:

Internet – a system that connects computer networks and allows for transfer of information.

World Wide Web (WWW) – the vast collection of Web pages available on computers connected to the Internet.

Web Page – a file on a computer that is viewable by a browser via the Internet.

Browser – computer program that makes it possible to view Web pages from other computers (e.g. Firefox, Internet Explorer).

URL (Uniform Resource Locator) – the address of a Web page or resource.

Domain – an extension code identifying the computer where a Web page is located (e.g. www.hpu.edu). Most common extensions:

- .com = commercial
- .edu = educational institution
- .gov = government (federal, state, or local)
- .mil = military
- .net = network (often personal Web pages)
- .org = (non-profit) organization

Search Directory – a site where links to Web pages are collected and arranged in a topic (or subject) format; like a table of contents (e.g. Yahoo, Librarian’s Index to the Internet).

Search Engine – software programs designed to locate information on the Web based on the user’s search words (e.g Google, AltaVista).

Meta-Search Engine – software programs that search multiple search engines (ProFusion, Dogpile).

[NOTE ON SEARCH ENGINES – each engine or meta-search engine only accesses part of what is available through the Web. Not all engines work the same way, so you must learn the effective search methods for each one. They do not evaluate the quality of what is found – the user must do that. The results of some search engines may be mostly commercial sites].
EASY TO USE --- HARD TO FIND

The Internet has something for everyone, but is that “something” enough to get the quality of information you need to produce a paper worthy of a good grade? Too many searchers start and END with the Internet. The material that is readily available on the Internet is labeled the “Visible Web.” This Web is estimated to contain only about 20% of the information that can be obtained via the Web.

An alternate source of information is the “Invisible Web” – a resource limited to selected searchers: through intranets, like Pipeline, or through purchase agreements, like HPU’s databases and eBooks, or through society memberships and other access points. This should become the major location for your information gathering, because these “invisible” sources will provide a higher quality of reliable material to support your paper’s thesis.

Although there has been much talk having full text books and journals available online and Google is working with major libraries to digitize older books, the reality is that most information produced is still only in print (probably about 80%). For this reason, your libraries will continue to be an important place for your information gathering process.

Most basic entry points to the Web (search directories or engines) make it very easy to start using the Web. Unfortunately, that doesn’t mean finding exactly what you need is equally easy. The Web does NOT use a standard vocabulary, like subject headings used in library catalogs or research databases, nor does it organized information by subject, credibility, usefulness, or any other criteria.

The different search engines cover varying Web pages and therefore do not produce identical results. No one site covers everything! An example of a comparison between retrieved results from Google and Yahoo can be found at the following site: http://www.langreiter.com/exec/yahoo-vs-google.html. The searched word, “synerge,” can be replaced by another word or words of your choice. Try it out!

In addition, how you create your search statement makes a BIG difference in the results you get: the words you chose to search by, how you combine those words, and even the order in which you type the words. Did you know that putting a phrase in quotation marks (e.g. “global warming”) keeps the two words beside each other (see next page)?

Another factor in the results of a search is the ranking process used by the search engines. Often the more scholarly resources will not fall within the top rankings of your retrieved results. See the section USING THE INTERNET FOR RESEARCH later in this guide.

If you want to learn more about using search engines, try these excellent links:

Search Engine Watch: Web Searching Tips http://searchenginewatch.com
Learn searching tips, search engine listings, ratings, statistics, and more.

InfoPeople: Best Search Tools http://www.infopeople.org/search/tools.html
Find links to the best search tools, and to their Help and advanced search options.

Plan your search: words and format

- What are the key topic words of your topic? In most search engines you do not need to type in a sentence -- just use the key words. This works like the Boolean AND search below. **radiation is a cause of cancer** (keywords are in bold)

- In some search engines and resource collections, if you put a string of words in their search box, it will look for every occurrence of each word -- and then dump them all together (an OR search below)! **That means EVERY occurrence of EACH of these words ("radiation" "cancer") will be retrieved, not just the results that bring these two words together.** [Use OR only for synonyms, when you want to broaden a search and find more results.]
• Many search engines now use the default “and” to connect your words. The Boolean connector **AND** allows you to search for results that contain all your words together in every site. If this method is not the default in your search engine you may need to type the **AND** between words. [Use **AND** to narrow a search and reduce your results.]

**BOOLEAN forms:**

![Diagram showing BOOLEAN forms]

**Basic formats in a Boolean arrangement:**

(some search engines will use symbols to replace the Boolean words – **AND, OR, AND NOT**)

- **AND** use as described above (ALL words will be in EVERY result)
  - cancer AND radiation
- **OR** use for synonyms or related terms
  - cancer OR neoplasm
- **AND NOT** use to exclude a term that you don’t want
  - cancer AND radiation AND NOT treatment

( ) use to “nest” items for special processing; usually for synonyms

( cancer OR neoplasm) AND radiation

**Additional search arrangements:**

- **“ ”** use for a phrase or proper name – this keeps the words beside each other
  - "greenhouse effect" "bill gates" "world bank"
- *** or ?** use for truncations and wildcards, to get variant spellings
  - advertis* for advertisement, advertising, advertiser
  - behavi?r for behavior and behaviour

**EVALUATION CRITERIA**

To determine if the information you have found is something that you would feel comfortable using in a research paper, you will need to evaluate the quality of the information and its source. These criteria should be applied to ANY information you find, but are particularly important for the Internet, where all levels and types of information becomes mixed together.

The criteria given here are built upon the **CRAAP Test**, which was developed by the Meriam Library at the California State University, Chico. The official CRAAP Test, retrieved 6/15/06, is available at [http://www.csuchico.edu/lins/handouts/evalsites.html](http://www.csuchico.edu/lins/handouts/evalsites.html)
Use these questions to evaluate any resource you plan to use in writing your research paper.

**Currency** – the timeliness of information
When was the information published? Has there been any revision to the material? A Web page should show when it was last updated.
Are the references as current as the publication?
Does the date of the publication fit the needs of your paper?
Web – are the links active?

**Relevance** – the importance of the information for your needs
Does the information relate to your topic or answer your question?
Is the information scholarly enough for your needs? Do you understand what is being said?
Have you compared this source with other sources in order to select the best one?
Is the page user-friendly and useful, with minimal advertising and with links to appropriate sites?

**Authority** – the source of information
Who is the author or publisher or sponsor? A Webmaster may not be the actual author of the information.
What are the credentials or organizational affiliation of the author?
Does the author have the qualifications needed to write on this topic? Does the topic fit with their training?
Is an email or contact address given for the author or publisher or organization?
On the Web the URL will give some information about the site (.com, .edu, .gov, .org). (“Internet Basics” lists domains – some are more reliable than others). Note – anyone can create an organization and get a .org domain address.

  REMINDER: what looks like a reliable source (.edu) may actually be someone’s personal opinion and not reflect the institution/organization. Look for the ~ symbol in the URL. (Example: hawaii.edu/~personalname.html).

**Accuracy** – the reliability, truthfulness, and correctness of the informational content
Where does the information come from? What was the original publication information, if it was first published elsewhere.
Is the information supported by evidence? Look for references identifying sources for factual information.
Was the information reviewed or refereed? Scholars go through this process before their articles are published in journals.
Is the information verifiable in another, credible, source?
Are there spelling, grammar, or other typographical errors?

**Purpose** – the reason the information exists
Is the material meant to inform, sell a product, persuade to an author’s opinion, entertain, other?
Do the authors/sponsors make their intentions clear?
Is the information fact (with references) or opinion or propaganda?
Is the writing unbiased and free of emotion or is it clearly slanted to a particular aim or audience?
Who is the intended audience (scholars, buyers, voters)?
How is the site presented? Is it well-organized? Are the graphics clear and helpful or distracting?

  NOTE: you can sometimes find more information about a source by slowly removing the elements between slashes, starting at the right end, until you reach the home page of the domain. Look for a link that says “home” or “about” or something similar.

Here are some further sites for critical thinking instruction (Note – they may use slightly different words or meanings than is used here in the CRAAP Test)

   Covers eight (instead of five) criteria, with questions and an example site to illustrate each criterion.
   Gives questions to ask and points to be aware of for five criteria, with several examples for each.

   Provides a tutorial on Web evaluation and five exercise questions.

Some of the illustrations presented in the above sites will be examples of “hoax” pages. There are many “hoax” pages on the Internet and some look very realistic. Therefore critical analysis is important to all users. Ultimately, you must be able to explain why you would use a chosen page for your research paper.

COMPARISON OF WHAT IS PUBLISHED

BOOKS
- Editors review and authors rewrite (maybe many times)
- Publishers must maintain a quality level to stay in business
- Libraries purchase, based on quality and reputation

ARTICLES
- Editors review and authors rewrite (maybe many times)
- For scholarly journals, the author’s peers review for quality
- Publishers must maintain a quality level to stay in business
- Indexing services select journals by quality and reputation
- Libraries purchase journals and indexes, based on quality and reputation

WEB PAGES
- No one is in charge or oversees what is uploaded
- Anyone can create or say anything on the Internet, whether it is true or not
- There are no standards to follow or reviews made
- Search engines have no selection by quality

You should expect that material obtained in or via a library has gone through various levels of review, whereas you are responsible for the critical review of Web resources.

NOTE: there is NO resource, in print or online, that will find ALL the research materials that have been produced, which is why you must use several different tools (library catalogs, periodical databases, and reliable search engines, like Google) to find the resources you need for your papers.

USING THE INTERNET FOR RESEARCH

HPU Library Pages are located on the main HPU page from the pull-down menu under Student Services

Log in to HPU Pipeline and use the Libraries tab for:
- Databases – a variety of general or subject specific databases, some with full-text.
- ULPL – University Libraries Periodical List – magazines, journals, and newspapers available at HPU.
- eBook Collections – online, full text books available through our subscriptions.
- Research Tools and Tutorials – frequently-used online links.
- Useful Internet Sites – selected subject links (why search the Web when we’ve already found a good site for you to use?).
- Links to other libraries.
- News and updates from the HPU Libraries.
- Links to services and forms.
- Library hours, maps, and more about HPU Libraries.
Library-selected e-resources are not the same as “online” resources. The former are expected to be reliable sources of information.