

PHIA  
Public Health Information Alliance  
<http://www.hslls.pitt.edu/phia>  
Barbara Folb, Project Manager  
412-246-5501  
[folb@pitt.edu](mailto:folb@pitt.edu)

## **Using the Internet Level 2 Locating Resources on the Internet**

# Table of Contents

Internet Resources Used in this Class .....	4
Search Tools.....	4
Compiled Resource Lists .....	4
Databases .....	4
Web Sites .....	5
Goals and Objectives .....	5
Using Search Engines .....	6
Steps in Searching.....	6
Think Before You Search .....	6
Example .....	6
Choose a Starting Point.....	7
Construct a Search Strategy .....	7
Basic Search Techniques .....	8
Use Specific Terms, Phrases, or Combinations of Each.....	8
Using Unique Terms/Phrases.....	8
Examples .....	8
Using Quotation Marks for Phrases .....	8
Using + and – .....	9
Examples .....	9
Using OR for Synonyms .....	9
Advanced Search Techniques/ Tools in Google .....	9
Google Advanced Search Features .....	10
Limit by Language, File Format, or Update Date.....	10
Limit by Location in Document (Occurrences).....	10
Page-Specific Searches, Similar Pages and Linking Pages .....	10
Search Within a Website (Domain) .....	10
Specialized Searches in Google .....	11
Froogle (Google Product Search) .....	11
Google Groups .....	11
Google News Search.....	12
Google Images Search .....	12
Advanced Search Techniques/ Tools in AllTheWeb .....	12
Meta Search Engines.....	12
Vivisimo.....	13
Surfwax.....	13
Search Engine Comparisons .....	14
Search Engine Showdown .....	14
Search Engine Watch.....	14
Using Compiled Resources .....	14
Cookies and Browser Behavior .....	15
Setting the Browser to Prompt for Cookie Acceptance .....	15
Characteristics of a Good Web Source .....	15
Evaluation Criteria .....	15
Further Information.....	15

Citing a Web Resource ..... 16

# Internet Resources Used in this Class

## ***Search Tools***

AllTheWeb

<http://www.alltheweb.com>

Google

<http://www.google.com>

Surfwax

<http://www.surfwax.com>

Vivisimo

<http://vivisimo.com/>

Search Engine Showdown

<http://www.searchengineshowdown.com/>

Search Engine Watch

<http://www.searchenginewatch.com/>

## ***Compiled Resource Lists***

APHA Public Health Links

[http://www.apha.org/public\\_health/](http://www.apha.org/public_health/)

The Grey Literature Page

<http://www.nyam.org/library/greylit/index.shtml>

MEDLINEplus

<http://www.medlineplus.gov>

Partners in Information Access

<http://phpartners.org/>

PHIA Internet Links

<http://www.hsls.pitt.edu/guides/phia/links>

## ***Databases***

CDC Wonder

<http://wonder.cdc.gov/>

ToxNet

<http://toxnet.nlm.nih.gov/index.html>

## **Web Sites**

CDC

<http://www.cdc.gov>

Evaluating Web Resources from Widener University, Wolfgram Memorial Library

<http://www2.widener.edu/Wolfgram-Memorial-Library/webevaluation/webeval.htm>

The Good, the Bad and the Ugly, or Why It's a Good Idea to Evaluate Web Sources

<http://lib.nmsu.edu/instruction/eval.html>

PHIA home page

<http://www.hslls.pitt.edu/phia>

Style Sheets for Citing Internet & Electronic Resources:

Humanities (MLA & Chicago), Scientific (APA & CBE), and History (Turabian)

<http://www.lib.berkeley.edu/TeachingLib/Guides/Internet/Style.html>

## **Goals and Objectives**

- Use basic and advanced search features of major search engines
- Use the major public health portals on the Internet
- Navigate the PHIA Internet Links Resource
- Evaluate the content quality of an Internet site
- Continue to build your skills after class using resources presented on the Internet

# Using Search Engines

## *Steps in Searching*

### Think Before You Search

Before opening your favorite search engine and beginning the search, a few moments of thought on the following questions will increase the chances of finding what you need. Clarifying what your topic is and what level of information you need will determine the best resources to use.

1. What is the question to be answered? Can you state it in specific terms?
2. Is your information likely to be free on the Internet? If not, you will do better to start with a source such as PubMed that indexes journal literature.
3. Do you want a general overview of a topic or a more specific piece of information? For a general overview you may want to start with a directory rather than a search engine.
4. Is there an organization or website whose information you trust likely to have that information on the web? If so, you may want to start there rather than with a search engine.
5. Is the information for a health professional or for the general public? If for the general public, does it need to be at a certain literacy level?
6. Is the information likely to be contained in a database that is not probed by search engines such as the AIDS Public Use Data in CDC Wonder?

### **Example**

You are working on a presentation on tobacco smoking in pregnancy, and tobacco cessation programs specifically for pregnant women. Think of the questions you want to answer in preparing for the talk. Where would you start looking for the answer to each question?

## Choose a Starting Point

Based on your initial evaluation of the question, choose a starting place. Here are a few types of resources and examples of each:

### **Single search engine**

A search engine is often your best starting point, unless you already know a site likely to contain the information. (ex: looking for a CDC guideline, start on the CDC's website)

AllTheWeb

Google

### **Meta search engine**

Use as a second line search tool. Searching multiple search engines at once, they pool results, but you lose the ability to use search tools specific to each search engine.

Surfwax

Vivisimo

*For more search engines and meta search engines, see Search Engine Watch, <http://www.searchenginewatch.com>*

### **Directory/Compiled Resource**

These vary widely in size and topic scope. Start here if you know of a good compiled source on your topic.

Open Directory

Partners in Information Access for the Public Health Workforce

Grey Literature Page from New York Academy of Medicine

Yahoo Directory

MEDLINEplus - Consumer Health Resource

### **Database not indexed by search engines**

It is extremely important to recognize that the information you want might not be indexed by search engines. This is true for resources that don't have a static page, but are generated as needed out of a database.

ToxNet – contains multiple toxicology related databases

CDC Wonder – statistical reports and datasets of interest to public health

FluNet – database of flu surveillance information

## Construct a Search Strategy

Your search strategy will depend on your starting point. Most search engines allow the use of the strategies illustrated in the following Basic Search Techniques section. If you are unsure how to construct a search in a particular search engine or database, consult the help section for that resource.

## ***Basic Search Techniques***

### **Use Specific Terms, Phrases, or Combinations of Each**

The more specific terms you use, the smaller your retrieval will be, and it will be more likely that pages on your topic will appear near the beginning of the retrieval results.

#### **Example**

Using Google, enter the following terms one at a time, searching following each added term. Note the type of pages retrieved for each search, and how many pages are retrieved:

smoking  
pregnancy  
statistics  
pennsylvania

### **Using Unique Terms/Phrases**

If you are looking for information likely to contain a unique phrase, use it as a first pass.

#### **Examples**

1. Use Google to find information on the EPA's ambient air monitoring reference by entering that phrase in the Google search box.
2. Use Google to find the National Vital Statistics System from the CDC.

### **Using Quotation Marks for Phrases**

Using quotation marks in most search engines forces the results to include only those pages where the phrase occurs exactly as entered. It also allows the searching of stop words. Stop words are common words usually ignored by a search engine, such as and, the, if, and so forth.

#### **Examples**

1. Put quotation marks around the phrase "ambient air monitoring reference" used in the last example. Note the decrease in retrieved pages.
2. Try the following strategy in Google and in AlltheWeb with and without the quotation marks: "hepatitis a" "food contamination" eliminate reduce

## Using + and –

- Using a + in front of a word or phrase indicates that it **MUST** appear in the results in the exact form used.
- Using a + turns off stemming, inclusion of variant forms of a word in the results, in Google. AlltheWeb does not appear to use stemming.
- Using a – in front of a word or phrase indicates that it **MUST NOT** appear in the results.
- - is very useful for removing unwanted concepts from the results, reducing retrieval.

### Examples

Search in Google for a word with and without the + and compare the retrieval for number of records.

Use Google to find pages on quarantine of people, not animals or imported goods. Start with just the word quarantine. Next, take identifying words from some of the off topic pages in the results, and add them with the -, as shown here:

quarantine –animal –plant

## Using OR for Synonyms

Using OR allows you to retrieve a concept that can be represented in more than one way. Before using OR, check a search engine's help page to learn the syntax for that search engine.

### Examples

Look for web pages discussing FlueNet and SARS.

Google search: flunet sars OR “severe acute respiratory syndrome”

AllTheWeb standard search: flunet (sars “severe acute respiratory syndrome”)

AllTheWeb Advanced Search, Boolean : flunet and (sars or “severe acute respiratory syndrome”)

## ***Advanced Search Techniques/ Tools in Google***

The Google and AlltheWeb Advanced search pages offer a number of tools that are useful for focusing a search. Some of the more useful ones will be reviewed here.

## Google Advanced Search Features

To access the following features click on the Advanced Search link from the Google home page. If you have an active search on the Google main page it automatically transfers you words into the Advanced Search page where you can further modify it.

### **Limit by Language, File Format, or Update Date**

Chose from the drop down boxes provided. Keep in mind that the update date on a web page may not reflect an actual date for the content.

#### **Examples**

- Find pages in Korean on tuberculosis updated in the last year.
- Find Powerpoint slides that mention radon abatement

### **Limit by Location in Document (Occurrences)**

Possible limits include title, text, url, and links. Limiting to title is a good way to find pages where your topic is the main point of the page.

#### **Example**

Find documents with the words “diesel retrofit” in the title of the document.

### **Page-Specific Searches, Similar Pages and Linking Pages**

When using these two features, type in the address of the page without the http:// prefix. Similar pages attempts to find pages that share characteristics with the page you specify. This tool is not always very useful. Links finds pages that link to that page. Because links may occur on pages generated from databases, it will not be a complete list, but it is often very useful.

#### **Example**

Find pages that link to the EPA voluntary diesel retrofit program page, <http://www.epa.gov/otaq/retrofit/>.

### **Search Within a Website (Domain)**

This is an extremely useful search technique. It can reduce some of the frustrations of finding information in a large site with an organizational scheme that is not ideal. Google and other search engines can search within a web site. This often provides a better result than using the search engine provided as part of the site because it will use the search engine’s ranking algorithm to bring the most on target hits to the top of the list. It also works better than searching the whole web for a topic in certain cases. Please note that portions of a website generated from a database will not be searched.

## **Example**

Do a search on Google of the whole web on the topic of influenza vaccination guidelines. Do the same search of the CDC website using the Google search within a site feature. Compare your search results.

Google search: influenza practice guidelines

Google search of the CDC site:

1. Go to the Google Advanced Search page. The link appears next to the search box on the Google main page.
2. If your previous search statement, influenza practice guidelines, does not appear in the search box labeled “with **all** of the words” type it in there.
3. In the box labeled “Domain” type in [www.cdc.gov](http://www.cdc.gov).
4. Click on the Google Search button.

## ***Specialized Searches in Google***

Google has an ever-evolving assortment of specialized search tools. Some appear as tabs on the main search page, others are accessed by clicking on the Services & Tools link on the main search page. New tools in a testing phase are available on the Google Labs page linked from the Services & Tools page.

## **Froogle (Google Product Search)**

Not just for personal shopping, this product makes it easier to find merchandise on the web. Like regular Google, there is a basic and an advanced search.

## **Example**

Find sources of sump pumps. Use both the basic and advanced search pages and compare results.

## **Google Groups**

I learned this trick from the library’s systems department. They use it to find advice on solving computer hardware and software problems. For just about any information topic, chances are someone has discussed it in a Usenet online forum. This tool allows you to find postings in a discussion group’s archives. On topics that raise controversy expect to find both reasonable and outrageous expression of opinion and fact.

## **Example**

Find postings on radon remediation.

## Google News Search

A very current search of multiple news outlets on the web, including television, newspapers. A newly added feature sends news alerts to your email on topics you define.

### Example

Search for “Allegheny County Health Department.”

## Google Images Search

Search for images based on keywords that appear attached to the pictures or on the same page. Before using any image in your own work, contact the page owner for copyright clearance. Google does not filter for public domain images.

### Example

Search for hand washing images. Try two different searches “hand washing poster” and “wash your hands.”

## ***Advanced Search Techniques/ Tools in AllTheWeb***

AllTheWeb has some of the same tools as Google, and some that Google doesn't have. Features found on AllTheWeb not found in Google include:

- Video, Audio, and ftp file search sections. Each of these sections includes a simple and advanced search screen.
- Complex Boolean search syntax. Uses four operators; and, or, andnot, rank; and allows parentheses. This is available on the Advanced Search page.

AllTheWeb and Google offer some of the same features with different parameters.

Examples include:

- AllTheWeb limit by date allows you to pick a date range using exact dates going back to January 1, 1980. Google goes back one year.
- AllTheWeb allows limiting of searches to specific parts of a page, such as the title, as drop down choices next to the search entry boxes on the advanced search page. Different choices can be made for each search box. Google has it as a choice further down the page that applies to all search words.

## ***Meta Search Engines***

A meta search engine doesn't maintain a database of web sites, it instead sends your search statement to multiple search sites at once. On the surface this sounds like a great time saver, but in practice users often sacrifice the ability to use the search syntax best understood by each individual search engine.

If you wish to use meta search engines, I recommend that you try Surfswax, <http://www.surfswax.com>, or Vivisimo, <http://www.vivisimo.com>. They each offer tools for sorting and managing search results.

## **Vivisimo**

Vivisimo attempts to organize the search results into clusters around topics. Relevance of the clusters presented varies from topic to topic.

### **Example**

Search for “syndromic surveillance”. Explore the organization of the clusters.

## **Surfswax**

Surfswax tried to make surfing your web results a smoother experience. Its SiteSnap tool allows you to see your search words in context in the web pages retrieved before going out to the website.

The Focus feature presents narrower and broader terms related to key terms it picks from your search statement and results, and lets you refine your search with them.

### **Example**

Search for asthma prevention programs

- Use the SiteSnap feature to review one of the sites retrieved.
- Examine the Focus words, and their narrower and broader terms. Find one that seems logical and use it to narrow the search results.
- Run the same search in a single search engine and see how the retrieved documents compare.

## ***Search Engine Comparisons***

Search engine features change frequently. To keep up with the latest on search engines, use the following resources:

### **Search Engine Showdown**

<http://www.searchengineshowdown.com/>

Compares and evaluates search engines from the searcher's perspective.

### **Search Engine Watch**

<http://www.searchenginewatch.com/>

While this resource is most interested in the search engine industry and the needs of web site developers who want to rank near the top in search results, it provides a great deal of information for the search engine user on search engine features and changes in a very timely manner.

## **Using Compiled Resources**

Directories and other resources compiled by hand vary widely in the quality of their content and organizational schemes. Usually the user has a choice of searching or browsing the resources by topic.

### **Examples**

Find information on hepatitis b for consumers on the MEDLINEplus website. Compare using the site's search feature to using the Health A to Z index.

Use the PHIA web links section to find information on tobacco control laws in Pennsylvania.

Use the Partners in Information Access website to locate information on grants and funding sources.

## **Cookies and Browser Behavior**

Cookies are small files stored on your computer by websites that you visit. Benign uses of cookies include:

- Storing your preferences for a particular website
- Recording your activity during a session so that search results can be recalled, reused or further manipulated

Cookies can be used to track your activities in ways you may not appreciate, such as causing advertising to appear based on your Internet use pattern. If you wish to have more control over the storage of cookies on your computer, you can change the settings of your web browser. For example, you can ask the computer to get permission from you before storing cookies on the computer. The downsides of setting for more control are:

- More pop up messages asking you to make decisions about accepting cookies.
- Potential for reduced performance in some websites.

### ***Setting the Browser to Prompt for Cookie Acceptance***

- Go to through the Internet Explorer menu choices Tools/Internet Options/Privacy/Advanced
- Put a check mark in front of “Override Automatic Cookie Handling.”
- Under First Party Cookies (cookies from the site you are currently viewing) and Third Party Cookies (from a website other than the one you are viewing, usually provides content such as ads) click in front of the word Prompt.
- Put a check in front of “Always allow Session Cookies”. Session cookies are temporary cookies that control the behavior of your current website visit, but are not stored on your computer after you leave.

There are other options for setting the browser to handle cookies. To learn more about cookies and how they change browser behavior, read the Internet Explorer Help.

## **Characteristics of a Good Web Source**

### ***Evaluation Criteria***

- a) Source reputation
- b) Easily accessible author information
- c) Objectivity
- d) Advertising/content separation
- e) Copyright info
- f) Last update info
- g) Depth of information presented

### ***Further Information***

For a more in-depth discussion, see the following:

Evaluating Web Resources from Widener University, Wolfgram Memorial Library, by  
Jan Alexander and Marsha Ann Tate

<http://www2.widener.edu/Wolfgram-Memorial-Library/webevaluation/webeval.htm>

The Good, the Bad and the Ugly, or Why It's a Good Idea to Evaluate Web Sources, by  
Susan E. Beck, Head, Reference & Research Services Department, New Mexico State  
University Library

<http://lib.nmsu.edu/instruction/eval.html>

## **Citing a Web Resource**

If you need to cite a web source, use the following web site as a guide:

Style Sheets for Citing Internet & Electronic Resources:

Humanities (MLA & Chicago), Scientific (APA & CBE), and History (Turabian)

<http://www.lib.berkeley.edu/TeachingLib/Guides/Internet/Style.html>