Workshop: Search the Internet
Compiled by Doris Grasserbauer – dgrasserbauer@ccny.cuny.edu

Topics:
1. General Information about the Internet
2. Different Search Engines and How to make your search as efficient as possible
3. Other Useful Information

1. General Information about the Internet

- Some Terms

- **Internet**: is a global network connecting millions of computers.

  The Internet has millions of users worldwide, and that number is growing rapidly. Almost all countries of the world are linked into exchanges of data, news and opinions. The Internet is decentralized by design. Each Internet computer, called a *host*, is independent. Its operators can choose which Internet services to use and which local services to make available to the global Internet community. Remarkably, this anarchy by design works exceedingly well. There are a variety of ways to access the Internet. Most online services, such as America Online or Verizon, offer access to some Internet services. It is also possible to gain access through other Internet Service Provider (ISP).

**Basic Internet Services** in chronological order as they got invented:
- **Email**: Short for *electronic mail*, the transmission of messages over communications networks.
- **Ftp**: Abbreviation of *File Transfer Protocol*, the protocol used on the Internet for sending files.
- **Telnet**: A terminal emulation program for TCP/IP (Transmission Control Protocol/Internet Protocol) networks such as the Internet. The Telnet program runs on your computer and connects your PC to a server on the network.
- **News**: An on-line discussion group. On the Internet, there are literally thousands of newsgroups covering every conceivable interest. To view and post messages to a newsgroup, you need a *newsreader*, a program that runs on your computer and connects you to a news server on the Internet. Both Microsoft Internet Explorer and Netscape Navigator come with newsreaders.
- **Archie**: A program that enables you to search for files anywhere on the Internet by filename.
- **Gopher**: A system that pre-dates the World Wide Web for organizing and displaying files on Internet servers. A Gopher server presents its contents as a hierarchically structured list of files. With the ascendance of the Web, most Gopher databases are being converted to Web sites, which can be more easily accessed via Web search engines.

- **URL (Uniform Resource Locator)**: is the global address of documents and other resources on the World Wide Web.

  The components of a URL are as follows:

  ![URL Diagram]
- **TCP/IP**: Abbreviation for *Transmission Control Protocol/Internet Protocol*, the suite of communications protocols used to connect hosts on the Internet. TCP/IP is the de facto standard for transmitting data over networks.

- **IP address**: is an identifier for a computer or device on a TCP/IP network.

Networks using the TCP/IP protocol route messages based on the IP address of the destination. The format of an IP address is a 32-bit numeric address written as four numbers separated by periods. Each number can be zero to 255. For example, 1.160.10.240 could be an IP address.

- **Domain name**: is a name that identifies one or more IP addresses.

For example, the domain name *microsoft.com* represents about a dozen IP addresses. Domain names are used in URLs to identify particular Web pages. For example, in the URL [http://www.pcwebopedia.com/index.html](http://www.pcwebopedia.com/index.html), the domain name is *pcwebopedia.com*. Every domain name has a suffix that indicates which top-level domain (TLD) it belongs to. There are only a limited number of such domains. For example:

- **gov** – US government organizations
- **edu** - Educational institutions
- **org** – mainly nonprofit organizations
- **mil** – US military organizations
- **com** - commercial companies
- **net** – mainly Internet resource company (such as an Internet Service Provider)
- **us, ca, th** – USA, Canada, Thailand, ... countries on this planet.

List of domain name suffixes: [http://www.computerhope.com/jargon/num/domains.htm](http://www.computerhope.com/jargon/num/domains.htm)

- **Search Engine**: is a program that searches documents for specified keywords and returns a list of the documents where the keywords were found.

Although *search engine* is really a general class of programs, the term is often used to specifically describe systems like ‘Google’, ‘Alta Vista’ and ‘Excite’ that enables users to search for documents on the World Wide Web. Typically, a search engine works by sending out a *spider* to fetch as many documents as possible. Another program, called an *indexer*, then reads these documents and creates an index based on the words contained in each document. Each search engine uses a proprietary algorithm to create its indices such that, ideally, only meaningful results are returned for each *query*. As this algorithm is different for each search engine, in common there will be a difference in the found websites.

- **Browser**: is a short cut for *Web browser*, a software application used to locate and display Web pages.

The two most popular browsers are ‘Mozilla Firefox’ and ‘Microsoft Internet Explorer’. Both of these are *graphical browsers*, which means that they can display graphics as well as text. In addition, most modern browsers can present multimedia information, including sound and video, though they require plug-ins for some formats.

For more information on computer and Internet technology terms please visit: (see next page)
2. **Different Search Engines and How to make your search as efficient as possible**

- **Different Search Engines**
  - Search engine with no advertisement: [www.google.com](http://www.google.com)
  - Search engines with specific topics:
    - Education: [www.worldwidelearn.com](http://www.worldwidelearn.com), [www.askeric.org](http://www.askeric.org), [ericae.net](http://ericae.net) (information about educational assessment, evaluation, and research information)
    - Search engine for children: [www.yahooligans.com](http://www.yahooligans.com), [www.kidsclick.com](http://www.kidsclick.com)
    - Engine which searches several other search engines at once: [www.dogpile.com](http://www.dogpile.com)
    - Search engine for articles: [www.findarticles.com](http://www.findarticles.com)

- **Hints**
  - Use very specific keywords to narrow the amount of Websites you will receive.
  - Rephrase your keywords if you are not getting the information you were looking for.
  - Use operators like ‘AND’ (respectively ‘+’), ‘OR’, ‘NOT’ (respectively ‘-‘).
  - Use the option to search on certain domain (e.g. in Google: keyword site:.edu; in Alta Vista: +domain:edu +keyword.)
  - Use more than one search engine to compensate for lags in coverage.
  - Use different search engines for different documents you want to find.
  - Examples:
    - Use **Google** to search PDF files or documents.
    - Use **Alta Vista** for images (.jpg, .gif) or sound (.MP3, .wav) files.
    - Use **metasearch** services like **Dogpile** to search several search engines at once.
    - Use **subject based** search engines when possible, such as **MusicSearch.com** or **HealthWeb**, or the educational search engines mentioned above.
    - Use **search engines of other countries** if you are looking for special information about this country: International Directory of Search Engines: (URL: [http://www.searchenginecolossus.com/](http://www.searchenginecolossus.com/).) The search Engine Colossus lists search engines that cover specific subject areas all over the world.
    - Use Web mega sites such as **Yahoo** that register and/or "catalog" new Web sites by their subject area and content.
    - Use **Find Articles** to search for articles.

You will find additional information at the following Website:

3. Other Useful Information

- **Resource Websites**
  - National Academy Press: www.nap.edu
  - National Association for the Education of Young Children: www.naeyc.org
  - International Reading Association: wwwира.org
  - Teaching with the Internet, Dr. Sheila Gersh: www.schoollink.org/twin
  - Tech Resource Website from the School of Education: http://www.ccny.cuny.edu/education/mmcenter/resources.html

- **Evaluating Websites**
  - **Brief overview:**

  - **More detailed information:**

- **Citation of Electronic Media**
  Regardless of format, however, authors using and citing Internet sources should observe the following two guidelines:

  1. Direct readers as closely as possible to the information being cited. Whenever possible, reference specific documents rather than home or menu pages.
  2. Provide addresses that work.

  Documents available via the Internet include all sorts of documents and information. At a minimum, a reference of an Internet source should provide a document title or description, a date (either the date of publication or update or the date of retrieval), and an address (in Internet terms, a uniform resource locator, or URL). Whenever possible, identify the authors of a document as well.

  More information:
  - **Commonly asked questions regarding how to cite electronic media:**

  - **Citation Machine:**

  - **Permission request template:**