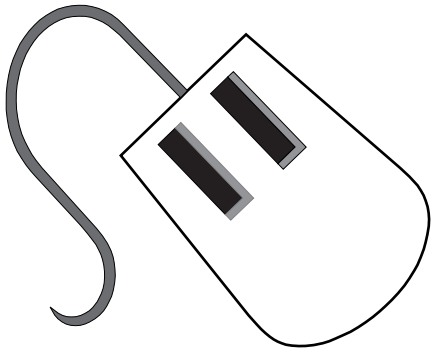


CHAPTER 7
Finding Information on the Internet

INTRODUCTION

Information overload is a fairly recent phenomenon. It is defined by one author as "the inability to extract needed knowledge from an immense quantity of information. . . ." (Nelson, 2001, para 7). Prior to the emergence of the Internet, most students had to visit a school or local library to find books and other resources that contained information being researched. Each student was limited to the amount of information that could be obtained by the number of resources in a particular library's collection and the willingness of the student to seek out further information from distant sources (e.g., via mail). Information overload was not a term associated with day-to-day research prior to the Internet. The emergence of the Internet has brought a whole world of information and disinformation to anyone with access to the Internet, but the transition has not been without new problems.

The term "Internet" is tossed around loosely these days but has come to mean the network of computers that are all connected together in the world for the purpose of sharing information and communicating with one another. The Internet includes the WWW (or Web), e-mail, Instant Messengers (IMs), FTP, Telnet and other related communication protocols. Most people are familiar with e-mail, IMs, and the Web. While e-mail and IMs are parts of the Internet, these two components of the Internet are more private and individualized (e.g., you do not receive every e-mail sent to every person in the world; typically, you should only get e-mails sent to you). The Web is the huge collection of webpages and websites that are usually available for anyone who is on the Internet and seeks them out. If you want to learn more about a particular breed of dog, you can search through tens of millions of webpages of information related to dogs to try and find what you are seeking. This chapter will help you to sift through the constantly growing Web to find information more efficiently and effectively.

Most people are coming to accept the Web as a powerful new resource in the quest for obtaining information, but the task becomes more and more difficult as webpages are being added to the Web every second of the day. One popular website that provides a historical timeline showing the growth and emergence of the Internet is called the Hobbes' Internet Timeline (Zakon, 2004, available online at *http://www.zakon.org/robert/internet/timeline/*). This site reports that there were more than 46 million web servers at the start

of 2004. A web server is a computer that is typically not used by people each day; rather, this computer is usually kept in a closet or somewhere safe and is left on all of the time to do its job of serving webpages to users. When another computer is instructed to access information (e.g., when you type a web address in your browser you are requesting that webpage from the server) then the web server receives the request for a particular webpage and grants permission and sends the webpage to the user's computer. Many servers host many different websites and many more webpages. So, there are many more webpages than the 46 million web servers. In fact, one popular **search engine** (a tool for searching the Internet) called Google recently reported that it was currently searching through well over 4 billion webpages (Google, 2004, available online at *http://www.google.com*). And, Google is only searching the pages that they have found as there are likely many more pages that have yet to be discovered by the Google company and many more pages being uncovered every day.

By any account, the Web is huge and contains more information than any other resource in the world . . . ever. These days the problem is not the difficulty of finding answers, but rather lies in sorting through a great deal of related or somewhat-related information to address specific information needs. In other words, information overload!

Finding What You Need—Fast!

Searching the Internet

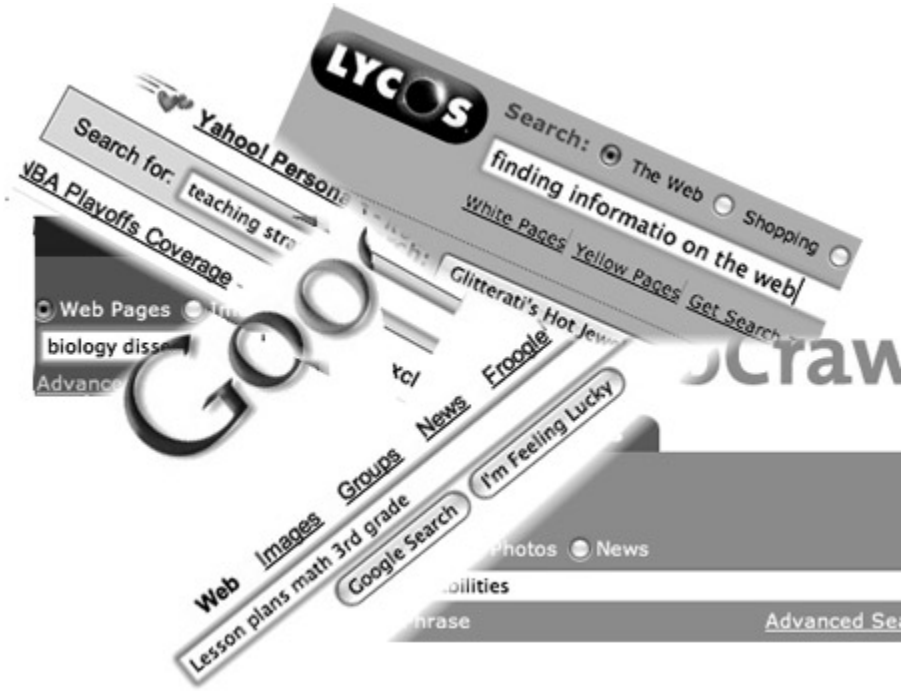
Information overload is a real problem that is only getting worse. Sifting through all available information during a specific Internet search can take considerable time and is not always a productive use of time. Learning to find quality information fast is a life skill with professional and personal benefits. Time spent improving Internet search skills will be recouped many times over. Searching the Internet is a dynamic process involving several steps or processes. It is not always a linear process—you may circle back and repeat steps as you refine your search. This chapter presents strategies for finding information on the Internet.

First, Define What You Are Looking For

Before trying to find information, some planning time for reflecting about the information you need can improve your final results. You might think about whether you want to get a general idea about something you know little about or perhaps you are looking for specific information about a narrow topic. If you are looking for people, addresses, types of files (such as audio) or other specialized items, you will want to use a specialized search tool, described later in this chapter. However, for text information, a search using keywords should suffice. If there are constraints, such as you only want hits from scholarly or professional teaching journals, then you need to keep these in mind when you choose your search tool.

Keywords

When doing an Internet search, you will often visit a search engine and type a word or series of words into the search box. These words are called keywords. Keyword selection



can make or break your search. Spending the time to think of various words that describe what you are seeking is time well spent. If you only have a broad concept or idea in mind (e.g., Michigan, U.S. Government, Civil War), you may want to start with a **subject directory** to help you find key words (such as Yahoo or About.com). A subject directory is different than a regular search engine in that humans create the subject directory by reviewing websites and placing these websites into various categories. This is a very time consuming process. As you navigate through a subject index, you can go from very general topics down to specific topics contained in the broad subject by clicking on links that help to narrow the topic. One lesser-known subject index is the Librarians' Index to the Internet (LII available online at: *http://lii.org/*). LII's motto is that they claim to provide "Information You Can Trust." While the advantage to a subject index is trustworthiness and accurate information, the disadvantage is that they tend to be significantly smaller than a regular search engine in the number of sites they contain because people have to review and classify each site contained in the database of sites for the subject directory.

Most keyword searches will return exact matches only (e.g., disability will not return results containing only the word disabled) so you may need to come up with synonyms or alternative terms and make those searches as well. Online dictionaries and thesauruses can help here (e.g., *http://www.m-w.com/*). Google provides a feature that allows you to search for synonyms automatically. You can use this feature by placing a tilde sign (~) immediately in front of the keyword you are

using. For example, if you did a search for "~sports" then Google would automatically search for sports as the keyword, but it would also likely search for athletics, play, recreation, etc. to help give you a more comprehensive search. A nice feature built into some search engines, like Google, is that the search engine will correct a misspelled word and will return the results for the misspelling, but will also recommend that you try the search again using the corrected spelling. This is a great benefit to students who are not great spellers, but also to those people who tend to make typos.

Most quests for information start with a search engine, a meta-search engine or a subject directory (also called a subject guide). In addition, there are portal sites, such as libraries or education portals that are often useful. Which you use depends on the type of information you need and how much you know already.

Individual Search Engines

No search engine indexes the entire web because it is impossible to keep up with the growth of the Internet and the billions and billions of pages contained on the Internet. Computer programs (called spiders or bots) create the indexes used in various search engines. These bots scour the Internet and visit pages to check for the title of the page, the text that appears on the screen, the keywords that are hidden in the code, and any other information that can be quickly ascertained. This information is all collected to create the index used by the search engine. Each search engine uses a different algorithm for helping the person doing the search to find the most relevant results. One search engine might try and return sites that many other websites are linking to because it figures these sites being linked to the most are popular and are more likely to be what the web surfer is seeking. Another search engine might place the most importance in the html meta-tags (these are hidden code statements that provide keywords), which are often specifically created to help search engines. And still another might place the most importance on the number of times a keyword appears on the webpage or where it is

The term, "Google" is actually a play on the term "Googol," which is the mathematical term for a 1 followed by 100 zeros. Google's play on the term reflects the company's mission to organize the immense amount of information available on the web. *http://www.google.com/corporate/index.html* Did you know that Google is much more than just a typical search engine? That's right, Google offers many other services beyond just searching the web, though the web-searching is their signature feature. Google has added an image searching resource (*images.google.com*) that provides a tool for people looking for specific images (e.g., to include in a PowerPoint presentation—assuming that you have copyright clearance on the image). They also provide features that you may not realize in the actual search box. For example, try typing 9+7*2 in the search box—that's right, the search box is also a calculator you can use. Now try entering your address—notice that it likely can return your name, phone number and a map to your house (assuming you have lived there for some time). Pretty amazing! To see a list of all of the nifty features available at Google, try this link: *http://www.google.com/options/index.html*

located on the webpage. The different and varied algorithms are what cause each search engine to return different results when the same search has been entered.

One way to get more familiar with making effective searches is to get to know two to three search engines and how they work. For example—do you need to put phrases in quotes? Can you truncate words? Do plus and minus signs have meaning? Boolean operators accepted? Do these terms even make sense to you? If not, then you likely have much to gain by reading the search tips on various search engine sites. If you examine the various search sites, you will notice that differences in the Search Language each accepts, which is why many people stick to one or two resources for their searches. There are real differences that can change the way you conduct more advanced searches from search engine to search engine, so this is something to explore as you try a new search site.

Some of the most well-known search engines are AltaVista, Google, Ask Jeeves, Hotbot, and Lycos. Google is the most widely used search engine on the Internet (see info box on Google). To use a simple search engine, you enter the search site's URL and when

Operator	Example	Read as . . .
AND	Woman AND female	Find all pages containing both the terms "woman" and "female" (narrows your search—pages must contain both terms)
OR	Woman OR female	Find all pages containing the term "woman" as well as all the pages containing the term "female" (broadens your search—pages can contain either term)
NOT	Kennedy NOT John	Find all pages that contain the term "Kennedy" but do not show the ones that contain the term "John"
NEAR	High NEAR School	Find all pages that contain the terms "high" and "school" within so many words or spaces of each other (varies by search engine)

the home page comes up you will see a box to type in your search term or key words. You should be able to easily narrow or broaden your search through an advanced search feature or use of Boolean terms (see box and table below).

Boolean Searching

Sometimes, you may have the need to search for a topic very specifically. If you wanted to look up "bass" to find information about the fish then you would likely start with "bass." When you type in "bass" as your keyword, you might notice a lot of the results are specific to bass guitars. A Boolean search uses special words written in capital letters to broaden or narrow your search. The most widely accepted Boolean search terms are "AND," "OR," and "NOT." So, you could try your search again using "bass NOT guitar" to give your search more specificity. Boolean searching only works if the search engine accepts Boolean words—Google, for example, does not accept all Boolean terms. You should be able to visit

the help webpages of your favorite search engine to see if they allow Boolean searching. You may even find that they allow this type of searching but have differing strategies for enabling this feature. The most common Boolean terms are shown in the table below. You might also note that multiple Boolean terms can be used in the same search.

Most search engines now recognize the space in between key words as meaning and/or so that they user does not even have to type "OR" or "AND" any more to get the same effect.

Ask Jeeves is different from many other search engines in that it allows the user to type in a search request using natural language and full sentence questions. They call this Natural Language Processing. They are probably still focusing on the keywords and discarding the irrelevant and common words, however (e.g., the, and, on, etc., get dropped).

Searching in the Classroom

Our society is increasingly becoming more reliant on being able to access information accurately and quickly. Most students have not learned effective ways to search the Internet and they will not be as competitive in future education and employment without these skills. In fact, many teachers do not know how to do effective searches so they will not be able to pass on search strategies to students, nor will these teachers be able to benefit greatly from the Internet as a teaching resource. Learning strategies for finding information should be a key objective for any school system in this technological age, yet this still appears to be an issue with little real attention. No standardized national or statewide tests focus on understanding how to access online information, so this trend is not likely to improve any time soon. Considerable time can be saved if students are aware of effective search strategies.

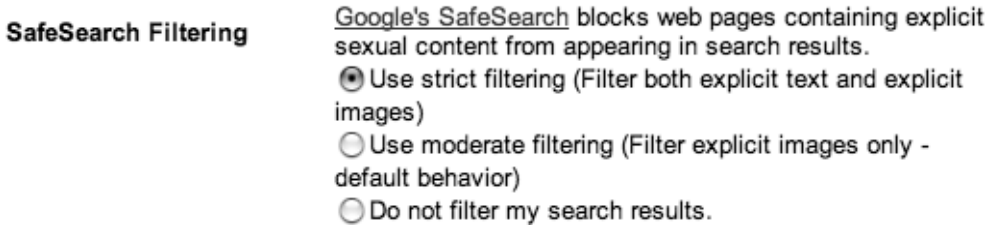
In the elementary school grades, students will probably not do a lot of searching on their own; however, there are search engines designed for younger students. Yahooligans (*http://yahooligans.yahoo.com*) is one example of a children's subject directory. This particular site has websites that have been carefully selected to ensure that the pages viewed will be age appropriate. As you teach a site like this to your students, you might provide a demonstration to help your students see how the site works. As you do the demonstration, you can really help your students to gain a better understanding by modeling the appropriate search strategies you are thinking about and using out loud. In fact, you should really think out loud so that your students can see the thought processes you use to best find the information you are seeking. Make sure you give your students time for practice and an opportunity to receive feedback from you. One strategy is to have your students complete a scavenger hunt (that you've prepared ahead of time).

Teaching Searching Example

Mrs. Brock teaches 4th grade. She wants her students to write a little report on space. She has a few resources in her classroom, but she also wants to use the lesson as an opportunity to teach some Internet search strategies. She schedules time in the school's computer lab and takes her class down to the lab. They use yahooligans. She has the whole class enter "space" in the search box and hit the space button.

The resulting page has a few different sections, but Mrs. Brock directs her students to the Category Matches section and tells them that this section is used to help narrow down the topic. Students who want to focus on specific components of space are directed to click on the link that most closely aligns to their topic of interest. For this example, Mrs. Brock tells her students that she wants to write about Mars, so she clicks on Astronomy and Space to find a whole host of new and more specific categories. She then goes on to show them how she narrows her search even more by selecting Solar System (294 related sites) and then Planets (79 related sites) and finally Mars (16 sites). She will use information found on these 16 pages for her class report.

As students move on into middle and high school, regular search engines (e.g., Google) can increasingly be used. Most schools should have filtering systems in place to help ensure that inappropriate sites are not found. You can also turn on safe searching from inside of the Google preferences.

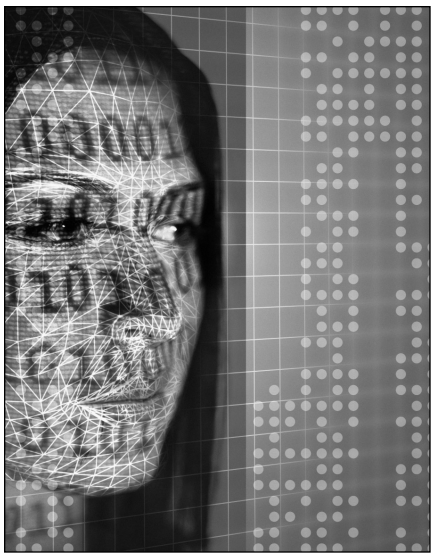


Older students can begin to learn about the tips and tricks of searching the Internet, which is why teachers should be learning these skills as well. Students can learn the nuances that various search engines and subject indexes employ and how to find these discrepancies prior to searching.

Even if the Internet is not the end goal of a particular search, the strategies used for searching the Internet can be applied to other similar situations. For example, many schools have a library. The library databases will often use similar search strategies and Boolean search operators as search engines, so the benefits of teaching these strategies can stretch far beyond the average classroom.

Summary

Information overload has become an issue that most schools are encountering today. Trying to make sense of the vast online world is a skill that is more and more important. Search engines are one tool that can be used to help provide some direction when seeking new information. One popular search engine is called Google, but many such tools exist. These search sites are typically distinguished by the algorithms they use to return relevant results to you.



To effectively use a search engine, a user would supply keywords and employ search strategies that are unique to a search site. These search strategies include the use of Boolean searching terms (e.g., AND, OR, NOT). Other search strategies are specific to various search sites (e.g., ~, "quotes", searching specific domains).

Teaching search strategies to students is becoming a necessity these days. Far too many schools are not realizing this need yet. Students can begin to learn search strategies as early as elementary school, and there are search sites that are developed with elementary school students in mind.

Discussion Questions

1. Why is it hard to determine how large the Internet is today?

2. What is the difference between a search engine and a subject index?

3. What are the most popular search engines available?

4. Why are Boolean search strategies effective?

5. Should Internet searching be something that is tested on Standardized tests? Why or why not?

Key Terms

Boolean searching: A searching strategy that allows the user to combine words and phrases by using certain words (e.g., AND, NOT, NEAR, OR).

Internet Filters: Software or hardware that is controlled by parents/teachers/authorities designed to prevent certain content from reaching the user (e.g., foul language).

Keywords: The specific words used to request information from a search engine.

Search Engine: A tool for searching the Internet (e.g., Google, Hotbot, MSN Search).

Subject Directory: A tool that is typically used by humans working to review websites and placing these websites into various categories.

Extension

1. Select 3 popular search engines and compare their special features. Do any offer more features than the others? Do they return similar results using the same search? Why do you think this is?
2. Google offers many other searching options beyond web searching. Please visit *http://www.google.com/options/index.html* and explain how two of these other options could be applied to your future classroom.
3. Are libraries becoming obsolete given the vast amount of resources available online? Explain your rationale.

References

Google (2004). Available online at: *http://www.google.com/*

Find It Fast: Avoiding Information Overload. Retrieved June 22, 2004 from: *http://www.sba.gov/test/wbc/docs/techno/finditfast.html*

Lyman, P. & Varian, H. (2000). How Much Information 2000? Retrieved March 1, 2004 from: *http://www.zakon.org/robert/internet/timeline/#Growth*

Nelson, M. (2001) We Have the Information You Want, But Getting It Will Cost You: Being Held Hostage by Information Overload, Retrieved March 1, 2004 from: *http://www.acm.org/crossroads/xrds1-1/mnelson.html*

O'Neill, E., Lavoie, B. & Bennett, R. (2003). Trends in the Evolution of the Public Web 1998–2002. Retrieved March 12, 2004 from: *http://wcp.oclc.org/*

Yahoo (2004). Available online at: *http://www.yahoo.com/*