Researching The Internet

Donald Ginesi
Marketing/Engineering
Consultant

March 2005

Finding Information on the Internet

What Search Tools Are Available?

Directories

One type of search tool is called a searchable subject directory. When you go to the home page of a directory you will generally find displayed a list of a dozen or more broad categories. Each category is broken down into additional levels. The intention is to browse through a directory in the same way you would find information in a filing cabinet filled with tabbed folders. Browsing through the levels of a classification hierarchy can help a researcher who is not sure on how to narrow a topic. Subject directories are good for general or popular topics, or to get product information. For example, assume you are trying to find computer sales in the Northeast. This would be difficult on a search engine because there are mostly common words in the search phrase. You might have better luck by using a search directory and starting with computers. The rule of thumb is the less specific your search words, the better the results.

Yahoo is a very well known subject directory. Let us assume we want to plan a trip to Las Vegas. You do this on Yahoo by clicking Yahoo Web Directory from the home page (www.yahoo.com), and then clicking on the following:

Travel/By Region/U.S./Nevada/Cities/Las Vegas

Closed directories, such as Yahoo, have dedicated editors to select and annotate links for every category. Open directories rely on volunteers to do this function. An example of an open directory is DMOZ (http://dmoz.org). I do not have a preference for open vs. closed directories and tend to use booth.

Directories generally offer a search function on key words that allow you to get results from deep in the directory without browsing. This is a search of the words in the links and annotations, not of the text. Because directories are small, typically a few million links, they may supplement results by simultaneously searching the key words on a general search engine, and adding these hits to the results page. Results from searching within the directory are generally labeled separately from the results using a search engine. A researcher uses a directory because of the selectivity of sites its offers, and not to search the web in general.

Search Engines

Search engines, such as Google (www.google.com), are organized using keywords or phrases and allow you to research large numbers of topics. They use algorithms to find relationships and compute correlations between subjects. Different search engines use different methods to determine these correlations, and you can learn how a specific search engine functions by going to sites such as www.searchengineshowdown.com. Search engines are generally better for specific or unusual topics, to see if any information exists at all on the topic, or if you are doing an exhaustive search. The rule of thumb with search engines is the reverse as for a directory; the more specific the

search words the better the results. There are no size restrictions to search engine data bases and they typically index billions of web pages. They capture the full text of all pages on a site, not just the title and information of the home page, like a directory. A search engine is preferred for exhaustive research on a topic because it will produce a much greater number of links than will a directory.

The Internet can be confusing in regards to classifying a search tool. For example, Yahoo started as a directory. Then Yahoo added search capabilities for the web, originally using Google's data base, to supplement the limited results from its own directory. Now Yahoo has purchased AllTheWeb and Altavista, two general search engines, and offers access to their combined data bases via the search box on Yahoo. While Yahoo originally started as a directory, it can also be called a search engine. This is true of many tools, and it is getting harder to classify them as one thing or another.

Google is generally accepted by many information experts as currently being the best search engine on the web (a brief guide to Google is included in this document). Google has the largest data base, contains excellent advanced searching techniques, and uses algorithms that frequently supply more relevant results to a given query. However, no single search tool, including Google, can give you all the information there is to be found on the Web for any given subject. It is estimated that Google contains less than 40% of all Web pages indexed within its data base. You will find that if you do the same query on two or more different search engines you will not see a large overlap in the results. In addition, while most informational professionals rank Google as number one, they also like to use other search engines that may offer advanced search tools Google does not. Never do just one search or use just one tool. You need to do the same search using different search tools for the best results.

Specialty Search Tools
Standard search engines perform best on more factual queries. They may not do so well on conceptual-type queries where you are looking for a complex answer with some type of analysis. A specialty search tool having a more focused data base may be a better option. Specialty search tools can be used to do many things, only some of which are:

- Locating people
- Human resource related issues
- Finding information on a competitor
- Finding information on a market or an industry
- Find information on various business and marketing topics
- Finding white papers and articles on scientific, academic or engineering topics
- Locating experts/consultants for hire
- Locating professional and trade organizations

The following is a list of specialty search tools:

Finding Information on the Internet

1. Targeted directory: composed by editors with specific expertise in one area. The Internet Intelligence Index (www.fuld.com/Tindex/I3.html) is an example of a targeted directory. It is designed to help you gather competitive intelligence information. It contains links to intelligence-related Internet sites, and is maintained by Fuld & Company, a leading firm in the field of Competitive Intelligence.

2. Focused search engine: a search engine that indexes web pages on specific topics or areas. LawCrawler (http://lawcrawler.lp.findlaw.com) is an example of a focused search engine. It is a legal search tool and data base that provides precision by enabling searches to be focused on sites and within specific domains known to feature high-quality legal information.

3. Portals: provide a start page for browsing designed to get you to relevant web sources quickly and easily. An example is www.msn.com. Portals are often used as a browser's start page, and are highly personalizable. They can contain all sorts of information on entertainment, sports, the stock market, weather, local interest stories, etc. They often will contain links to popular commercial and business sites as well. A portal will generally provide a search box on the home page to help find information on the Internet. However, I do not use portals for conducting advanced searches, only to quickly scan for news and other information when I first go on-line, or to find specific sites for which I can't remember the URL. Portals do not typically have the features required to do an exhaustive or advanced search. Yahoo is used by many people as a portal, even though it is a directory/search engine, because it offers the same features on its homepage as does portals like MSN. The selection of a portal is up to the personal likes/dislikes of the individual, and there are many to choose from.

4. Meta Search Engine: a search tool that can send queries to multiple general search engines and directories simultaneously. MetaCrawler (www.metacrawler.com) is an example of this kind of tool. MetaCrawler sends your query to most of the major search engines and directories, and then combines the results in a single list, eliminating duplicates. A Meta search engine is a good tool because it allows you to see how different search engines do against one another for a specific search phrase, and it can be very quick when you are doing a limited search. However, Meta search engines in general do not have the features for refining and doing an exhaustive search. You generally get fewer, and more broad results than you would with a standard search engine. In addition, Meta search engines do not return enough information to make a considered decision of whether to view a page. Titles alone do not suffice.

5. Value Added Search Services: are tools that combine a general search engine with a proprietary collection of fee-based material. Examples are NorthernLight (www.northernlight.com) and Edgar-Online (www.edgar-online.com).

6. Client based search tools: reside on your computer and not on the Internet. They tend to be more intensive in processing search queries. An example is Copernic (www.copernic.com).

What Search Tools Should I Use?

There is no definitive answer to this question. You will most likely use a combination of search tools, depending upon the subject being researched, the type of information required, the search strategy, how exhaustive you want the search to be, and your familiarity/preferences for specific tools.

The Invisible Web

The invisible web is a term used to describe web pages that are not, for various reasons, indexed by the crawlers of search engines. Invisible web content can include:

*   Disconnected pages
*   Pages composed mostly of music, video or images
*   Content in relational data bases
*   Real-time data
*   Dynamically generated content
*   Pages consisting of PDF, Postscript, Flash, Shockwave, programs or compressed files

The invisible web can be broken down into the following categories:

o   *Opaque web*: files that can be, but are not included in search engine indices.

o   *Private web*: pages that are capable of being indexed, but have been deliberately excluded from engines.

o   *Proprietary web*: pages only available to those who have special terms of exchange with the provider.

o   *Truly Invisible web*: pages that can not be indexed due to truly technical reasons.

Invisible web resources tend to be more focused than general search engines and directories, and provide better results for many needs. Besides more comprehensive results, invisible web sources offer more control over search input/output, increased precision and recall, and of course, information not available via standard search tools. You can look to the invisible web when you are looking for a particular answer, if you want to find exhaustive results (including those from the visible web), and when timeliness of information is an issue. Searching the invisible web may be useful when you are looking for information on: public company fillings, patents, library catalogs,

company directories, searchable subject bibliographies, and economic information and job postings.

Just as with the visible web, new Invisible web resources are being made available all the time. www.invisible-web.com is a website created by the author of a book of the same name. It is a good site to learn how to find information on the Invisible web. The Librarians' Index to the Internet (www.lii.org) is a searchable directory of web resources organized into categories including "best of", "directories", "data bases", and "specific resources". Most of the Invisible web content reviewed by LII falls in the "data bases" and "specific resources" categories. Free Pint (www.freepint.co.uk) is an e-mail newsletter dedicated to helping you find reliable web sites and search the internet more efficiently. The feature article covers a specific topic in detail and offers links to sites connected to the article's subject area. Quite often these are Invisible web resources.

As you become more familiar with what is available on the Invisible web you will need to build your own collection of resources. A good way to become familiar with the Invisible web resources is to explore the various gateways, cherry-picking resources that are most relevant to your needs. Organize your collection of resources for easy access on future searches.

A Brief Guide to Google
What's on the Google Home Page?
Of course there is the keyword input box. Alongside are links to more advanced search functionality:

Advanced Search: Google's default search allows you to do quite a bit. But Google's advanced search provides more options. Refer to the later section for more details.

Preferences: provides a nice, easy way to set your searching preferences. These include interface language (in which tips and messages are displayed), search language (acceptable languages for web pages being searched), SafeSearch filtering (avoiding results that may be offensive), number of results per page (10, 20, 30, 50,100), and whether you want results to open in the current window or a new browser window. For research it is best to have as many results on a page as possible.

Language Tools: Google offers several language tools, including one for translation and one for Google's interface.

Along the top, there are tabs linking beyond the default Web search to additional services:

Images: offers an archive of over 400 million images culled from sites all over the Web.
Groups: Usenet is a global network of discussion groups. Google has archived Usenet discussions for over 20 years.

News: Google continuously mines over 4,500 news sources for breaking news and the freshest of events.

Froogle: Froogle is a combination of the words Google and frugal and is a searchable shopping index that can get you right to the on-line point of purchase for an item you are interested in.

Under the input box are two buttons. The Google search button will give you a list of results from which to choose based on the keywords you entered. The "I'm Feeling Lucky" button delivers results based on what Google believes is the most relevant page given your search. For example, inputting "vortex shedding" as my key words gives a list of 55,300 items with the search button. The "I'M Feeling Lucky" button gives a Cambridge University Press- Gallery of Fluid Mechanics page describing Karmen vortex streets (the first result of the search list).

Asking For What You Want
The 10-Word Limit
Google has a limit of 10 words, keywords and special operators (described below) combined. Anything else will be ignored. Think carefully when you compose your search phrase. Limit your query to more obscure (or technical or jargon) words and phrases.

Phrase Searches
Google defaults to searching for pages with your keywords anywhere on the page, whether side-by-side or scattered through out. Enclosing the words in quotations is called a phrase search and will result in a list of hits where the key words only appear side-by-side.

Basic Boolean
How Google searches for the key words you enter is controlled by the Boolean operators you include in your search phrase. "Boolean" is just a fancy term to describe the words AND, OR and NOT when used as operators in a search phrase (these words are capitalized when used as operators). Google's default Boolean operator is AND. That means if you enter key words without any Boolean modifiers Google will search for

all of the words. For example, entering <u>Lunar Solar Eclipse</u> will return pages with the words lunar, solar and eclipse, although the words can appear in any order.

The Boolean operator OR will retrieve sources if either term appears. So if you enter <u>Lunar OR Solar Eclipse</u> you will get pages with the words lunar and eclipse or solar and eclipse.

The Boolean operator NOT will limit the responses to an inquiry. For example, if you enter <u>Lunar NOT Solar Eclipse</u> you will get pages with the words lunar and eclipse, but not pages with the words solar and eclipse, or pages that contain the words lunar, solar and eclipse. Google does not use the NOT operator.

Negation

If you want to specify that a search term item not appear in the results you must include a minus sign (-) in the phase, instead of the Boolean NOT operator. For example <u>windows software -glass</u> will give you results containing the words windows and software but not the word glass. The minus sign must appear directly before the word or phrase you do not want (no space in between).

Explicit Inclusion

In general, Google will search for all the words or phrases you enter.   However, there are certain words Google will ignore. "I", "a", "the", and "of" are some of the words, called stop words, Google will ignore. You can force Google to take into account a stop word in your phrase by putting a + in front of it. Stop words that appear inside phrases are not ignored and do not require the + operator.

Case Sensitivity

Google is case insensitive. However the Boolean operators must be capitalized.

Wild Cards

Stemming is adding a wildcard character (*) to part of your search terms. This will yield hits containing any variation of the term. For example, "<u>moon*</u>" would yield results containing moon, moons, moonlight, moonshot, etc.   Google does not support stemming. But Google does offer a full word wildcard where the wildcard substitutes for one full word in the phrase. For example, entering "<u>three * mice</u>" would yield results containing three blind mice, three blue mice, etc. One * stands for one word, two *s signify two words. Wild cards are useful for avoiding the 10 word limit because Google does not count wildcards toward the total. They are also useful for filling in the blanks on a fitful memory where you only remember some words of a key title or phrase you want to search.

Advanced Operators

Google offers some special syntax to help you refine your searches and get rid of meaningless results. Special syntax, or advanced operators, allows Google to search specific parts of web pages for specific types of information.

Finding Information on the Internet

"**Intitle**:" restricts your search to the titles of web pages. An example would be intitle: "green bay" which will yield results having green bay in the title of the web page.

"**Intext**:" searches only the body text and ignores link text, URLs and titles. This operator is good for search words that are too common in URLs and link titles.

"**Inanchor**:" searches for text in a page's link anchor, the descriptive text of a link.

"**Site**:" allows you to narrow your search by either a site or a top level domain. For example, lunar eclipse site:edu would limit the responses to those having the key words from only academic sites.

"**Inurl**:" restricts your search to the URLs of web pages. Using "inurl:" instead of "site:" has the advantage of being able to search subdirectories.

"**Link**:" returns a list of pages linking to the specified URL. Entering link:www.monitortech.com will get a list of links to the Monitor home page.

"**Filetype**:" searches the suffixes or filename extensions. Google indexes several different Microsoft formats, including PowerPoint (ppt), Excel (xls) and Word (doc). You can even limit your search to hits in PDF format (pdf). For example, "RF capacitance level" filetype: pdf would yield only PDF documents featuring the key word phrase.

"**Related**:" finds pages that are related a specified page. For example, related:www.google.com will return hits on a variety of other search engines.

"**Info**": provides a page of links to more information about a specified URL. The information includes a link to the URLs cache, a list of pages that link to the URL, pages that are related to the URL, and pages that contain the URL.

Advanced Search

The advanced search function has a separate page with "fill-in-the-blank" searching options. You can achieve most of the same functionality using the standard search with the Boolean and syntax operators just reviewed. However, you may find it easier using the advanced search method.
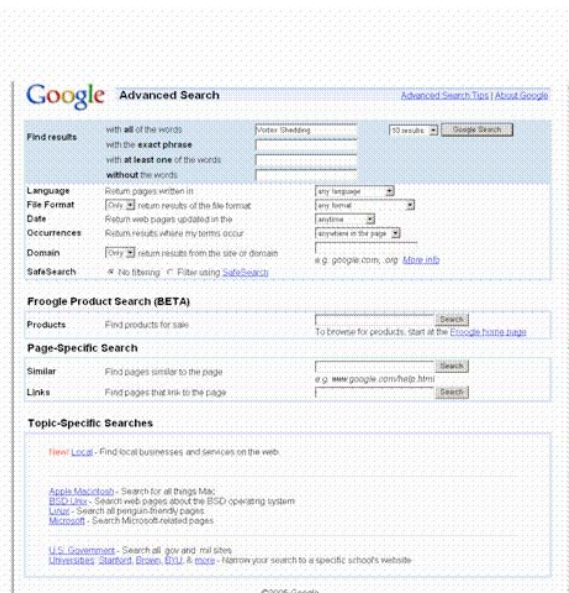
Finding Information on the Internet

Using the text boxes at the top of the Advanced Search page allows you to establish nuances of just the search phrase you are looking for. There are four boxes:

- With all the words
- With the exact phrase
- With at least one of the words
- Without the words

To the right of the word boxes is the results drop down list. By default, Google displays 10 results per page.

You can use the "Language" drop down list to specify what language all returned pages must be in.

You can use the "File Format" drop down list to include or exclude several different file formats from your search (like the "filetype:" syntax discussed earlier). The advanced search only includes a subset of some of the formats you can select. To get to the others you need to use the "filetype:" syntax.

The "Date" drop down list allows you to specify search results updated in the last three months, six months or year.

The "Occurrences" drop down list allows you to specify where the search terms should be found. Again this is similar to some of the syntax operators reviewed previously.

The "Domain" feature is an interface to the "site:" syntax. It also allows Google to not return results from a site or domain.

Search Results

Google encompasses so many search properties, and has so much data at its disposal that it fills every results page to the brim.
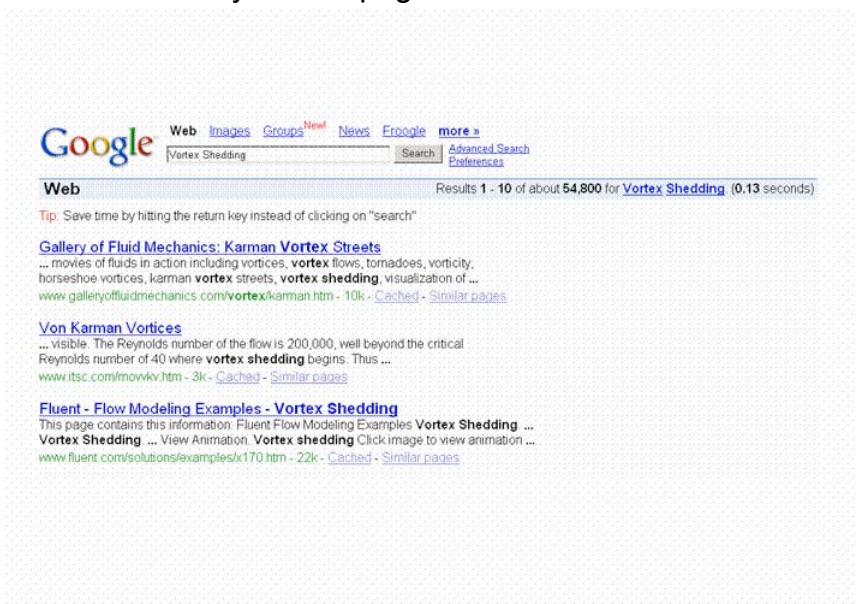
First, at the top of the results page is a selection of tables allowing you to repeat your search across other Google search categories besides web pages (Google Groups, Google Images and Google Directory). Beneath that you will see a count for the number of results and how long the search took.

Finding Information on the Internet

Sometimes you will see results/sites called out on colored backgrounds at the top or the right of the results page (unfortunately there are none in this sample). These are sponsored links.
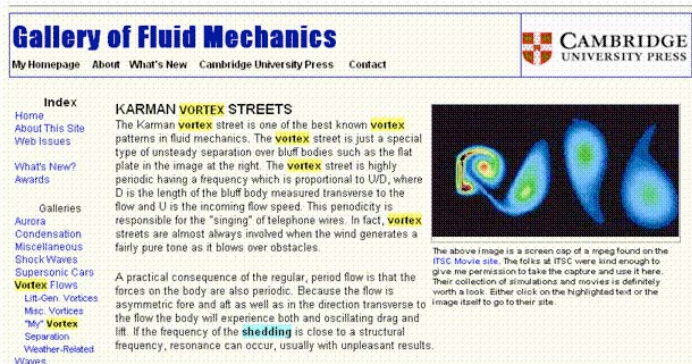
Beneath the sponsored links you may sometimes see a category list. You will see a category list only if you're searching for a very general term and your search consists of one word. When you see categories, you are seeing the information from the Google Directory. The Google Directory does not index all web pages, but describes whole sites instead (about 1.5 million URLs). This can make it a better search tool for general topics.

Most people will look only at the first screen or two of results. Search engines normally apply a mathematical algorithm to all hits arising from the search terms, giving each a relevancy score. Those with the highest scores will appear at the top of the list. The algorithm takes into account how often the word or phrase appears on a page, if the word(s) are near the top or title of the page, how close the words are to each other, and how rare a word is. Google has added some new twists to their ranking algorithm that regularly produces more relevant sites on the first and second pages than many other search engines. That is one reason it is preferred by many information professionals.

The top line of each result is the page title, hyperlinked to the original page. The second line offers a brief extract from this site. Sometimes this is a description or a sentence, sometimes it is HTML mush. The next line sports several informative bits of metadata (the URL, the size of the page). There's a link to a cached version of the page if one is available, and finally there is a link to find similar pages. The page size will only be available if the page has been cached.

Clicking on the cached hyperlink shows you a snapshot of the page with the search words highlighted (as shown on the right). This is a quick way to see if a questionable result is relative to your inquiry.

If you have a broadband connection the metadata is really not that important. But it can save you time with slower internet connections. Check out

the page summary and see where your keyword(s) appear. It may indicate the context is not what you are looking for. Check the size of the page (if available). Does the size of the page seem reasonable for the type of information you are looking for?
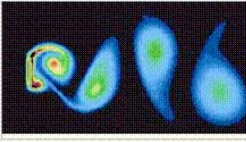
How to Develop a Search Strategy

The following template is one that you can use to retrieve the most useful and relevant pages on the Web:

1. *Prepare*
2. *Create*
3. *Run*
4. *Scan*
5. *Refine*

1. Prepare

The more preparation you can do before you begin a search, the better will be the outcome. Use the following checklist to define your inquiry prior to developing a search strategy:

- What is the topic you are searching?
- Do you know of any recognized experts related to the topic?
- What kind of information do you need: statistics, full-text articles, reports, special facts, public records, etc.?
- What do you think are the best sources of the information: a trade association, an industry discussion group, an academic center, a company website, a government agency, etc.?
- What do you already know about the topic?
- Where is the biggest source of the information you are looking for likely to be?
- How far back do you need to research?

This information will provide ideas on what kind of search tools to use, as well as ideas on your search statement.

2. Create

You can only get good results if your search statement is well done. Your keywords and phrases should be clear and unambiguous and contain unique words when possible. Of course the words you use should be contained in the text on the pages you are looking for.

As an example, let us assume you are trying to find out if the town square of Vilnius, the capital city of Lithuania, has a statue of Frank Zappa. First apply the check list shown above (under "1. Prepare"). The information desired is any web page or article or news item that refers to a statue of Frank Zappa in Vilnius. A standard search engine, in this case Google, will be the initial search tool. Vilnius, in addition to being the capital of Lithuania, is also the name of a university, and can also be someone's name. The name Frank Zappa is fairly unique. Construct a search phrase based on this

information. Vilnius is a clean and unambiguous noun, but the phrase "city of Vilnius" must be used to minimize false hits. Entering "<u>City of Vilnius" "Frank Zappa</u>" yields several results, the third of which informs us there is a bronze statue of Frank Zappa by the river, but not in the town square. Ambiguous words like situated, town, square, etc., were not needed and would not have added any value to the inquiry.

As another example, assume you wanted to find information from Life magazine (assuming Life magazine articles are on-line and free). Since the name Life is a common term, you would search using the phrase "<u>life magazine</u>". You can also try using the advanced "inurl:" operator. The search phrase becomes <u>inurl:www.life.com</u>. You can also restrict the search by using the "filetype:" advanced operator. The search phrase becomes <u>inurl:www.life.com filetype:pdf</u>.

Start with a very specific search query, and then get more and more general as you proceed with subsequent searches. You may become overwhelmed with results initially if you start off with a very general search phrase. Remember that you want to be more general if using a directory than if using a search engine.

3. Run
It is good to set a time limit for the amount of time you should spend doing a search on any single search tool. A reasonable time limit is 30 minutes. If you have not found relevant information in that amount of time then it is best to re-evaluate your strategy or try another search tool.

4. Scan
Once the search tool has provided a list of hits based on your initial search statement, you need to scan and analyze the results. Look at the total number of hits your search has created. Examine the first 100 results. You want to search the results carefully. Observe which sites have the most relevance to your topic. Review these sites and see how your key words appear in the title and text. Look for any other unique words or phrases that appear in these pages or for the names of authorities and experts in the field, or for any associated topics. Also scan the non-relevant sites and use the cached links to see how your key words appear on these pages. Also take note if relevant results are most often in one particular format (such as PDF), or in a certain particular site.

5. Refine
Once you have scanned your initial results you need to refine your search. Start by changing the order of the key words you have, and/or by adding or subtracting words. Rerun the search and see how the number of total hits and the number of relevant hits changed from the previous search. Try other unique words and phrases, as well as the advanced search features, to further refine your inquiry. A search engine will generally start ranking results that keep re-appearing in search after search higher on subsequent results lists.

As an example, assume you are trying to find how new sources of energy are being developed to fight global warming and reduce green house gases. During your searches you may have found that the most relevant sites use the term "climate change" frequently in reference to global warming. You may also have discovered "alternate fuels" is a better way of describing new sources of energy. Lastly, you may have noticed that the most relevant results seem to be in the form of PDF reports. You can now refine your search to read:
 "climate change" "alternate fuels" filetype:pdf.

When I first started using the Internet for research I would just jump in, and not take the time to develop a search strategy. For example, if I wanted to find papers on flow measurement, I would just go to Google, type in flow measurement, and see what results came back. I generally got thousands of hits, the vast majority of which were irrelevant. If I had not been lucky enough to find a valid result here and there from amongst the junk, I would generally get frustrated and give up.

Today, I am more successful and efficient in finding information on the Web by taking the time up front to develop a search strategy.

Tips to Assist You in Searching the Web

➢ Use more than one search tool for extensive searches. Stick to a few search tools that work best for you. Knowing how to use them will save time and yield better results.

➢ Many search tools will provide specialty toolbars that you can download and use to optimize the effectiveness of that tool. These specialty toolbars, such as the one offered by Google.com, integrate directly with your browser. You can do a search to see if there are any specialty toolbars associated with the subject you are trying to research using a general search engine. Enter a keyword(s) and the word "toolbar".

➢ Rare or unusual words are easier to find on web pages than common words. Try searching with the rare terms first in your string. For example, if you are looking for information on the Dow Corning Company you will get fewer misses if you use corning dow. Be sure not to reverse word order if you are phrase searching. You would enter "dow corning" and not "corning dow" because the words must both appear in the text and in that specific order when enclosed by quotation marks.

➢ Find and use synonyms for your most important concept in the search phrase. You can try doing a search with a keyword(s) and the word "means" to get different angles on the meaning of the word and potential pointers for doing your search. For example, suppose you were trying to find information on the medical condition "hydocephalus", and were not familiar with it, and therefore did not know of any synonyms. If you did a search on "hydrocephalus means" your hits

Finding Information on the Internet

might turn up phrases like "water on the brain", "spina bifida", and "neurology". You can also use words like "glossary" and "dictionary" to do the same thing.

➢ The more kinds of proper nouns you use to describe an item, the more your search results will be slanted towards it. For example, if you are trying to find information on "hawks", you can search using words describing different kinds of hawks, like "red tailed", "red shouldered", "sharp skinned", etc. Try running the search several times with various combinations of words or phrases to see what you get. If you don't know of such examples up front you can turn to Google.com for help. Sometimes you can get a list of related items to a key word(s) by going to http://labs.google.com/sets and entering the information you have available.

➢ If a search term includes special or non-alphabet characters try incorporating it into a search phrase using quotation marks.

➢ Use a Meta search engine to get a quick overview of what might become a very extensive search. You can also use a Meta search engine as an initial test for your proposed search phrase to see what kind of results it generates.

➢ Try using a directory to find words that you can use in a search phrase to submit to other search tools. There are also search tools that have features to help you create and refine search phrases.

➢ If you run a search and get back a very large number of results remember to try the following to screen out (or filter) unrelated pages on subsequent searches:
  • scan the initial results for new words and phrases;
  • use phrase searching;
  • use "filetype:" and other advanced operators.

➢ If you know the name of experts or prominent people related to a topic you can use them as filters to reduce the number non-related results. For example, Walt Boyce, Dick Miller and Wade Matter are three experts who have published a good deal of information on vortex shedding flow meters. If you were interested in searching that topic you could enter the search phrase: "vortex shedding" "walt boyce" OR "dick miller" OR "wade matter". If this exceeds the maximum number of words allowable for the search phrase you can break it up into searches for each individual.

➢ You can use the phrases "white papers" and "press releases" in your search statement to filter out non-related pages. White papers are generally free and contain detailed data and analysis on a subject. Since many white papers are presented in PDF format you could further filter the results by using the "filetype:" operator. An example would be "level measurement" "white papers" filetype: pdf.

➢ You can also use publications as filters, if they are relevant to the subject of the search and if the publication itself or some of its articles are found on the web.

14

Finding Information on the Internet

➢ Once you have accessed a web page, click "Edit" and use the "Find on this Page" feature to quickly locate a particular term you are looking for.

➢ Look for a site index and search box on a site. They are usually more useful for navigation than by means of graphics and links on a home page

➢ You can use a search engine to find specialized directories by searching for a subject and adding the word "resources" to the phrase. You can also find portals by adding the phrase "portal". Add the word "pathfinder" to get a site that provides a list of printed resources for the desired subject. Many libraries provide pathfinders that are guides to literature and resources in the library.

➢ www.extremesearcher.com is a website created by the author of a book of the same name, and is a good site to find valuable information on searching the internet.

➢ Use http://searchenginewatch.com to find out about the capabilities of individual search engines and comparison information. The site is maintained by Danny Sullivan, a leading journalist in the area of web search engines.

➢ Groups, newsgroups, and mailing lists are interactive forums that are often under-used by researchers. These tools can be goldmines of information. These channels allow you to take advantage of the knowledge and expertise of people you may not have in your Rolodex. The difference between these tools is in the method specific information gets to you. For groups, messages are posted on the Internet for anyone to read. Mailing list information goes to the emails of those who have registered for the list.

➢ You can use good search tools like Google.com to help search for information on specialty tools that do not have good internal searching capability. For example, Tile.net is a useful site to find small mailing lists or lists hosted by a university or small organization. However it is hard to search because it does not have a good search mechanism, and it offers little in the way of descriptions. Suppose you wanted to find mailing lists on "surfing". You could go to Google.com and enter surfing site:tile.net and get pointers to lists. Always try using the internal search function on a site before using another tool like Google.com because Google.com may not have all the pages from the site indexed. You can also use Google.com to find links and directories by using the words "links" and "directory" with your key word(s).

➢ Weblogs are sites that maintained by individuals, almost a form of personal journalism, and dedicated to a specific area or topic. Some weblogs, especially those concerned with searching the internet, are considered main line sources of information. Weblogs may not exist for your subject of interest. But it may be worthwhile checking to see by using sites like Feedster.com. They may be a good way to find the name of leading people with expertise on that subject.

Finding Information on the Internet
© 2005 Monitor Technologies, LLC.

➢ It is a good idea to copy and paste the top results of each search you run into a text file for reference later on. Any category of research can develop dozens or hundreds of lines of inquiry. You will be able to refer to previous search attempts if you take the time to organize the materials you find along the way.

➢ Listed below is an easy way to locate state, county and city government sites:

State:	www.???.state.??.us	(e.g. www.dot.state.oh.us)
County:	www.co.???.??.us	(e.g. www.co.fairfax.va.us)
City:	www.ci.???.??.us	(e.g. www.ci.milwaukee.wi.us)

??? stands for the agency, county or city name respectively, and ?? stands for the state.

➢ The best source for finding up-to-date press releases is the company website. However, not all companies make their press releases available on-line. The next step is to go to on-line trade and industry magazines and search their data bases on the company name for press releases past and present. You can also find several websites, like Thomas.net, that act as clearinghouses for press releases. You can search their data base by company name and/or products.

➢ The U.S. Patent Office has a searchable data base on patents. You can also check the status of a patent application if you know the inventor's name and/or you have the assigned number. It is difficult for the layman to do patent searches, and the best method may be to use a fee-based service (like GetThePatent.com). However, if you are game, there are a number of free websites, in addition to the Patent Office site, (like PatentCafe.com) from which you can search.

➢ There are many websites that offer "keeping up tools". For example, KeepMedia (www.keepmedia.com) is a fee based data base for magazine and print articles that will allow you to enter key words describing a specific topic. KeepMedia will periodically run a search against those key words and automatically e-mail you with any recent results. Such websites can be used to keep up-to-date on patent information, magazine articles, press releases, new products, web page content changes, companies, and general industry news.

➢ There are many law oriented websites, such as LawCrawler.com, that will allow you to search for and track current litigation in U.S. courts.

➢ Use the link feature on Google to find pages that link to a particular website. For example: link:www.monitortech.com will yield all the pages that link to the Monitor website.

Finding Information on the Internet

- ➢ On certain specialty search engines (like BetterWhoIs.com) you can do a "who is" search. You can enter a URL and find out who the site is registered to, as well as administrative and contact information.

- ➢ Before you start searching for industry information you need to find the SIC (Standard Industrial Classification) code, if you don't have it already. SIC codes are used to identify industries in census data, financial data, market reports and other useful data bases. NAICS (North American Industry Classification System) codes are in the process of replacing SIC codes. You can go to www.census.gov or www.osha.gov or other websites to find SIC/NAICS codes by industry.

- ➢ Dunn & Bradstreet, The Gale Group, and Plunkett Research print a number of resources for industry groups and specific industries which can be obtained on-line for a fee.

- ➢ The following texts are potential sources for industry information:

  - • Statistical Abstract of the US
  - • U.S. Market Trends & Forecast
  - • U.S. Industry and Trade Outlook
  - • Standard & Poor's Industry Survey

  These texts may be available at local public or college libraries. You might try large on-line libraries, such as the Library of Congress, to see if these books are available via the Web. You can also try Amazon.com, which now provides full text searches on-line for tens of thousands of books.

- ➢ Canada offers industry information on-line at Strategis (www.strategis.ic.gc.ca).

- ➢ The U.S. Department of Commerce Market Research Library allows you to search for international market research reports by industry.

- ➢ Some industries have extensive portals to industry and regulatory information. For example, www.foodprocessing.com and www.powderandbulk.com are websites for the food processing and materials handling industries respectively.

- ➢ The Census Bureau can be a good source for industrial reports. Industry trade associations and groups providing coverage of specific industries may have broader and more current information available than the Census Bureau. There are several websites, like www.galegroup.com , that can be used to locate trade and industry organizations.

Finding Information on the Internet

➢ The annual reports of companies within a specific industry can provide a great deal of industry information as well as company information. They also identify government regulations that affect the industry, and market share data. Annual reports should be available from the company's website. If not, several websites, like www.annualreportservice.com , can provide annual reports for many companies for a fee.

➢ You can find articles on various industries via on-line magazines. If it is a large industry then you can search the data bases of national publications like Business Week and Time. However, for many of the industries Monitor deals with you will be more successful with on-line industry and trade magazines like GrainNet.com, PowderandBulkSolids.com, etc.   BizJournals (www.bizjournals.com) is a major publisher of local and regional business journals and has an on-line data base of articles.

➢ A strategy to help ensure you are tapping into the most complete collections of government information is to use all the "one stop" clearinghouses that help users to navigate to various agencies and find what they need. Examples of such sites are FedWorld (www.fedworld.gov) and FirstGov (www.firstgov.gov).

➢ If you are looking for statistics you can try this approach. List search terms in a string. For example, if you were looking for energy statistics you could try words like "energy", "facts", "statistics", and "numbers".

➢ You can use data bases to find information on the web. There are thousands of searchable data bases on-line. Some consist of newspaper and/or trade journal articles, patents, market research reports, etc. Use your key word together with the words "database", "repository", "archive", etc. in a query to search for data bases. For example: "widget manufacturing" 1998 Archive.

➢ Many data bases are hidden in the reaches of larger sites. Go to a large site, such as the Library of Congress, and use the site search tools to look for data bases.

➢ There may be industry/market specific newsletters and mailing lists that will provide useful information. You can use websites like www.mediafind.com to look for relevant newsletters, and www.topica.com for mailings lists.

➢ There are many industry experts and consultants that you can contact to learn about specific industry trends and other things. You can search for such contacts using the web by going to sites like www.about.com .

➢ Trade associations can be a good source for industry information; however you may need to become a member to have access to the data.

➢ When all else fails you can always go into the wallet and purchase detailed market reports from companies like ARC Advisory Group and Frost & Sullivan. If you go to their websites you should be able to see what reports are available and browse abstracts and/or the table of contents. MindBranch (www.mindbranch.com) MarketSearch (www.marketsearch.com) and other sites offer market reports from multiple sources.

Useful Websites for Searching the Internet

This section contains websites that will help you find publicly available information on a wide range of topics. If these sites do not provide the information you are looking for directly, they may point you to secondary sites that do. There is a brief description of each website under the domain address. If you take a few moments to read through these descriptions you can get a good idea of the different information you can find on the Web.

Multiple website options are provided for most topics. This is because no one website may contain all available information. In addition, each website has a different "look and feel" and operation. Options allow you to look at different sites for retrieving specific information and choose the one you find most convenient to use.

The websites in each section are listed alphabetically. **The same website may be referenced in more than one section if so applicable**. Unless stated as being fee based, the websites listed herein are generally free. However, even on free sites you may need to register before you can have access to any information.

➢ Don't lose your library card. Local and university libraries often can provide access to many fee-based data bases at no charge or based on a minimal hourly rate.

The websites are categorized as:

General Search Tools
There are search engines listed in this section which have broad data bases and features to conduct inquiries on almost any subject. Google is the preferred search engine by many experts because it has the largest data base, and many excellent features for refining a search. There are sites that provide links to more search tools if you do not find what you are looking for in this list. There are also websites that will provide reference information on the various search engines listed, or on the various aspects of conducting a search.

Specialized Tools
This section contains search tools that have data bases that support specific topics. You can use these specialized tools to:

Finding Information on the Internet

- ❖ Find experts and expert advise
- ❖ Search the Web by asking questions and not using key words
- ❖ Do a "who is" search on a domain name
- ❖ Tap into academic and scientific data bases
- ❖ Find information on the invisible web
- ❖ Find industry newsletters
- ❖ Find specific search directories and other tools
- ❖ Find mailing lists
- ❖ Find industry discussion groups

There are also websites for Meta search engines.

United States Government
The governments in the U.S. – federal, state and local – tend to collect and disseminate more information than you can imagine, and also tend to use the Web to make this information available. The U.S. government is also the biggest wholesaler of private information. Federal and state government agencies routinely sell data bases to marketing companies and other interested buyers. This information can relate to your customers, your markets, or the industries you serve, and can include a significant amount of statistical data.

Almost every federal government agency is on-line. There are almost 5,000 government sites from more than 42 departments and agencies.

Because there are so many sites, you may want to turn to sites called government gateways. These sites organize and link government sites, in order to find the starting point for your research. Some gateways are simply collections of links.

This section contains direct links to various agency sites of interest to Monitor,   and government gateway sites.

News/Magazines/Articles/Technical Publications
This section contains a list of websites that will allow you locate the following:

- ❖ News articles/news papers
- ❖ Articles and papers from technical journals/proceedings
- ❖ Magazine articles/on-line magazines
- ❖ Books
- ❖ Educational resources
- ❖ Information on scientific/engineering topics

Business Tools
Information is the life blood of business and the Internet is plentiful with business resources. Websites allow information to be retrieved pertaining to:

- ❖ Available consultants/experts on a given topic/industry

❖ Company annual reports
❖ Marketing reports
❖ Specific industries
❖ Business/marketing reports on various subjects
❖ Trade shows

Keeping up Tools
This section contains a list of websites that will automatically keep you up to date, usually via email, on a variety of subjects. These subjects include web traffic patterns, news/business information, magazine articles, patents, industry news, etc.

Competitive Intelligence/Researchers
A list of websites for finding CI consultants/software.

Patents/Trademarks
A list of websites that allow you to track patent applications and awarded patents, as well as registered trade marks and copy rights.

Professional/Industry/Organization Locations
This section contains a list of websites that allow you to find experts on a given topic or in a given industry; search organizational data bases (such as the ISA, ISO and ASME) for standards and other documents, and locate trade/industrial/professional organizations.

Marketing Reports
This section contains websites that link you to available marketing reports or to marketing consulting firms.

Law/Legal Tools
This section lists three websites that can be used to track current litigation (e.g. on a competitor) in the U.S. courts.

Other
This section is a catch-all for websites that may be useful for collecting a wide range of information such as:

❖ quotations,
❖ best practices,
❖ FAQs,
❖ Trivia,
❖ Maps,
❖ Free software downloads,
❖ Tutorials,
❖ References, etc.

Also listed are the websites for organizations that control the World Wide Web (e.g. protocols and standards, domain names, etc.)

General Search Tools

www.alltheweb.com
www.altavista.com
www.yahoo.com
All three now use Yahoo's database. They offer Boolean search, include a cashe, have options for limiting a search, and index many types of files.

www.allsearchengines.com
You can use this site to link too many kinds of search tools.

www.calvin.edu/library/searreso/internet/as/
Alpha Search provides access to the finest Internet gateway sites.

http://dmoz.org
DMOZ is an open directory. The search function searches the entire directory or a category. DOZ has a unique descriptions box that will provide a "scope note" defining the kinds of things placed in a category. DMOZ search offers sophisticated search techniques including Boolean characters, phrase searching, the ability to identify specific words in the page title, and stemming (*).

www.extremesearcher.com
This is a website created by the author of a book of the same name. It is a good site to find valuable information on searching the internet.

www.gigablast.com
Fast search engine. Consistently retrieves relevant material. GigaBits ideas for related concepts and narrower searches are displayed at the top of the screen after a search.

www.google.com
Google is arguably the best and largest of the search engines. Currently offers the largest data base and the greatest amount of features.

www.hotbot.com
General search engine; also offers search engine for news groups. Easy to use with advanced search options.

www.infopeople.org/search/tools.html
From CA state library system – provides easy access to many of the best search tools.

www.lycos.com
Advanced search mode has many features geared to research. Drop down menu filters limit searches by words, URL and languages.

Finding Information on the Internet

http://marylaine.com/neatnew.html
Use this site to find the latest new sites for research on the web.

www.researchbuzz.com
Site dedicated to giving you information on search tools. Weekly newsletter keeps you up to speed various search engines and invisible web resources.

www.resourceshelf.com
ResourceShelf is created by Gary Price, who is an expert on finding information contained in the "invisible" web.

http://scout.cs.wisc.edu/scout/report/current/
Scout report provides summaries of the most valuable and authoritative web resources. Scout Report Signpost provides full text search of 6,000 of these summaries.

www.searchengineshowdown.com
Compares and evaluates search engines; provides information on how to develop effective search strategies.

http://searchenginewatch.com
Links to search engine reviews, search strategies and rankings; free and fee based newsletters available.

www.teoma.com
Teoma is an excellent search engine that points you to expert links or super sites on your subject. Owned and operated by Ask Jeeves. Teoma has "refine" and "resources" features. Refine offers ideas on how to narrow a search. Resources offer links to recommended sites created by subject specialists.

www.virtualchase.com
VirtualChase offers up to date information on search terms and online research techniques.

www.websearch.about.com
This site contains innovative and useful search techniques and research sites.

www.wisenut.com
WiseNut is the second largest database behind Google. Customized folders called "wise guides" help group results. Does phrase searching

Finding Information on the Internet

Specialized Tools

www.about.com
www.allexperts.com
These sites allow you to find experts and expert advice on a wide range of topics including management, marketing, agriculture, composites/plastics, and many industries.

www.askjeeves.com
Search engines that let you ask specific questions rather than query by key words.
www.betterwhois.com
www.register.com
www.samspade.com
These sites allow you to conduct a "who is" search. You can input a domain name and find information about the owner.

www.copernic.com
Site has free and fee based software that once downloaded allows you to search multiple engines simultaneously. It also allows you to create reports from the results and email these reports.

www.csa.com
Fee based website offers access to more than 70 scientific databases.

www.daypop.com
Current events search engine that crawls the web at least once a day. Allows you to enter a search subject and see which weblogs appear most often.

www.delphiforums.com
www.ezboard.com
These two sites provide means to locate Internet groups dedicated to specific areas of interest.

www.delphion.com
Delphion Intellectual Property network provides free access to a wide variety of data collections and patent information.

www.dialog.com
 Fee based site providing on-line based information.

www.dialogweb.com
Fee based search tool geared to scientific information.

www.dogpile.com
Meta search engine that includes About, Ask Jeeves, FAST, FindWhat, Google, LookSmart and Overture.

Finding Information on the Internet

www.factiva.com
Fee based tool geared for business information.

www.feedster.com
This site allows you to find any weblogs that might line up with the topic or area of your search.

www.freepint.com/gary/direct.htm
This site offers links to search tools to help you find "invisible" web resources.
www.globalspec.com
www.thomasregister.com
Sites allow you to search for manufacturers via product queries; links to manufacturers home pages; some biographical info on manufacturers. GlobalSpec allows you to search for technical papers and press releases by subject and offers a special toolbar you can download for free.

http://infomine.ucr.edu
Site covers data bases, directories, on-line library catalogs, e-journals and books. You can conduct a keyword search on many fields as well as by broad category. Focus on revealing academic-oriented "invisible" web documents.

www.ingenta.com
Searchable data base provides summaries for articles from more than 26,000 magazines and journals (mostly academic). It also provides an article delivery service (fee based).

www.intelliseek.com
Fee based; tap 300 search engines and 600 databases to find, analyze, filter, report and track information.

www.invisible-web.net
Offers links to specific search tools to help you find "invisible" web resources not indexed by traditional search engines.

www.ixquick.com
Meta tool ties into 14 search tools and translates your query into each engine's preferred syntax.

www.lexisnexis.com
Fee based tool geared to news/topical information.

www.libdex.com
This site contains a directory of library catalogs.

Finding Information on the Internet

http://lii.org
Serves as a directory to websites and identifies and describes searchable data bases. Precise search capabilities via its "advanced search" feature. Good site for finding invisible web resources.

www.mediafinder.com
Fee based archive of newsletter listings and national directory of magazines/catalogs.

www.metacrawler.com
MetaCrawler contains 11 search tools (including Google and AltaVista). It combines results in one list with duplicates deleted.
www.meta-list.net/query?acc=110en/
Site allows you to search for info in more than 170,000 public internet newsletters and email discussion lists.

www.moreover.com
One of the best search tools for news.

www.newsindex.com
Search engine for news information.

www.nexis.com/research
Pay service offering news and business publications.

www.northernlilght.com
Once a great free search engine, it is now fee based.

www.payperclicksearchengines.com
Site to find information about paid positioning on search engines.

www.profusion.com
Profusion combines the capability of 9 search engines and directories. Runs check to verify results are still live sites. Profusion also provides direct access to the Invisible web with the ability to search over 1,000 targeted sources (like the NY Times).

www.researchville.com
Meta tool allow you to target reference sites, dictionaries, encyclopedias, etc.

www.searchability.com
Site contains a collection of specialized search sites to lead you to other directories and tools.

http://searchpdf.adobe.com
This search engine provides only PDF files that correspond to your inquiry. This can also be done on standard search engines, like Google.

Finding Information on the Internet
© 2005 Monitor Technologies, LLC.

www.tile.net
Site you can use to locate internet discussion groups relevant to your interest area.

www.thrall.org/proteus.html
Easy to use Meta tool that offers to show you how to maximize results.

www.topica.com
Topic is a comprehensive searchable directory of more than 100,000 mailing lists to companies, associations, and individuals. Many associations are using Topica rather than managing their own lists.

http://vivisimo.com
Meta tool sorts results into folders.

www.vlib.org
Web directory that contains an excellent selection of specialized directories by category.

Government Sites
www.access.gpo.gov
This is a service of the GPO that provides free access to a wealth of important information produced by the Federal Government.

www.carrollpub.com
Site allows you to search Federal, state and municipal government directories.

www.ccr.gov
Data base of vendors that do business with the U.S. government. Search and obtain data on over 300,000 companies that are registered with the CCS.

www.census.gov
U.S. governments top site. You can find statistical data for any community or geographic region; contains historical data on various subjects; pay per print change on reports.

www.census.gov/cir/www
This site has data on production and shipment of selected products, focusing on the manufacturing of commodities, such as chemicals and industrial equipment.

www.census.gov/epcd/www/naics.html
NAICS has replaced SIC codes; site has downloadable table of codes, a search engine, mapping of old SIC codes to new NAICS codes

www.census.gov/statab/www/stateabs.html
Create maps using data from Census CD-ROMs; also features statistical abstract for the U.S.

Finding Information on the Internet

www.epa.gov/
This is the website of Environmental Protection Agency.

http://factfinder.census.gov
This site provides quick way to find a fact from the huge Census database.

www.fedstats.gov
This site provides access to official statistics collected and published by more than 70 Federal agencies. You don not have to know in advance which agency produces them.

www.fedworld.gov/
FedWorld is a website that offers a searchable directory of US government sites.

www.firstgov.gov
Clearinghouse of government sites linked to every federal and state agency. Easily locate business regulations and vital records.

www.gov/main/www/srchtool.html
This is a search tool for finding information on government sites.

www.gpoaccess.gov/index.html
GILS is an effort to identify, locate and describe publicly available Federal information resources.

www.irs.gov/businesses/small/article/0,,id=108149,00.html
IRS prepares reports to bring their examiners up to speed when auditing certain industries. These reports contain industry buzz words and general layout information.

www.ita.doc.gov
This site from the International Trade Administration provides industry statistics and data. It also offers a guide to resources such as key trade associations, market research companies and trade publications.

www.itds.treas.gov/commodityprof.html
Site assists businesses in locating industry data from over 100 federal agencies.

www.lib.lsu.edu/gov/fedgov.html
www.lib.umich.edu/govdocs/index.html
All 3 sites are run by universities and are tools to find federal information on government sites. University of Michigan site serves as clearinghouse for local, state, federal and international information. Each provides a comprehensive searching aid for government information on the web.

www.loc.gov
Library of Congress site; offers huge database of books, papers and images.

Finding Information on the Internet
© 2005 Monitor Technologies, LLC.

www.netronline.com/public_records.htm
Site has links to public records offices for state and local governments.

www.nist.gov/
This is the website for the National Institute of Science and Technology.

www.ntis.gov
Site to find federal government reports related to technology and science, including unclassified technical reports of government sponsored research.

www.odci.gov/cia/publications/factbook/index.html
This site contains the CIA's publications about international statistics and background information on other countries.

http://oregonstate.edu/dept/pol_sci/govinfo.htm
This site provides access to many U.S. Government databases in economics, education and demographics. See the home page of the GISP for a complete listing.

www.osha.gov/oshstats
This site allows you to search by company for any OSHA violations.

www.osti.gov
This site provides a portal to information about Federal research projects.

www.sba.gov
This site provides statistics about recently approved loans to small U.S. businesses.

www.searchgov.com
Site contains links to all federal web sites.

www.sec.gov
Good starting place for information about specific companies and industry trends. SEC requires all publicly held companies and some large private companies to report all their financial information.

www.stat-usa.gov
Site contains a subscription service by Department of Commerce that provides country profiles, commercial guides, international market research and financial statistics.

http://tech-net.sba.gov
This is an electronic gateway of technology information and resources for and about small high-tech businesses. It is a search engine for researchers and scientists, a marketing tool for small firms, and links to investment opportunities and sources of capital.

Finding Information on the Internet
© 2005 Monitor Technologies, LLC.

www.10Kwizard.com
Interface to SEC's EDGAR database. It provides free access and search to the EDGAR system. You can view current and historical filings, from the start date of each company's existence by keywords, phrases, and names.

http://thomas.loc.gov
This is a service of the Library of Congress. Thomas provides access to U.S. Government legislative information.

News/Magazines/Articles/Technical Publications/Books
www.academicinfo.net
Education focused subject directory for colleges.

www.awesomelibrary.org
This site contains a collection of educational resources.

www.bizjournals.com
This site is a major publisher of local/regional business journals whose articles can be found on-line.

www.bulk-online.com
This site serves as an Information center for powder and bulk solids handling and processing community.

www.chemicalequipment.com
This is the webpage for chemical equipment magazine.

www.chemicalprocessing.com
This is the webpage for chemical processing magazine.

www.chicagotribune.com
www.nytimes.com
These papers have good websites and allow you to search and retrieve articles from their archives. NY times requires you to register but is free (it is a part of the invisible web because spiders can not get past registration page into archive).

www.clpgh.org/
This is the home page for the Carnegie Public Library in Pittsburgh.

www.controleng.com
This is the webpage for control engineering magazine.

www.controlmagazine.com
This is the webpage for control magazine.

Finding Information on the Internet

www.csa.com
Fee based website offers access to more than 70 scientific databases.

www.daypop.com
Current events search engine that crawls the web at least once a day.

www.ei.org/eicorp/
Owned by Elsevier, world's leading publisher of technical documents; fees for purchasing documents and software.

www.feedandgrain.com
This is the webpage for Feed & Grain magazine.

www.findarticles.com/pi/index.html
Archive of hundreds of magazine articles since 1998.

www.foodprocessing.com
This site serves as the webpage for the food processing industry.

www.grainnet.com
Webpage for gain journal with news and information on the grain, milling, feed and seed industries.

http://groups.google.com
Site offers a free tool that catalogs the newsgroup portion of the Internet.

www.keepmedia.com
Fee based searchable data base for magazine and print articles. You can set up key terms and be notified by email of any hits. You can search the data base for free and see what is available.

www.iee.org/publish/
Fee based data base that provides coverage of the literature in engineering and sciences.

http://infomine.ucr.edu
Terrific site for academic and scholarly resources; searched by keyword and category. It focuses on revealing academic-oriented "invisible" web documents.

www.ingenta.com
Ingenta contains a data base that provides summaries from more than 26,000 magazine and journal articles (mostly academic). It provides an article delivery service (fee based).

Finding Information on the Internet
© 2005 Monitor Technologies, LLC.

www.iac-insite.com
Fee based service with a free trial; offers ability to search more than 1000 trade and industry publications, newswires, and business magazines on dozens of industries.

www.isa.org/journals/intech
This is the webpage for Intech magazine containing searchable archive of articles.

www.lexisnexis.com
Fee based data base geared to news/topical information.

www.libdex.com
This site contains a directory of library catalogs.
www.lii.org
Good tool when you are looking for anything academic or research oriented. Serves as a directory to websites and identifies and describes searchable data bases.

www.loc.gov
Library of Congress site; offers huge database of books, papers and images.

http://magportal.com
Conduct searches on the on-line version of over 450 magazines and web-only publications.

www.mediafinder.com
Fee based archive of newsletter listings and national directory of magazines/catalogs.

www.moreover.com
One of the best search tools for news.

www.newequipmentdigest.com
Webpage for new equipment digest magazine.

www.newsindex.com
Search engine for news information.

http://nvl.nist.gov/
 Searchable data base for the National Institute of Science and Technology.

www.nypl.org
This is the home page for the New York City Public Library.

www.powderandbulk.com
This site serves as the website for the materials handling professional.

www.powderandbulksolids.com
This is the webpage for power & Bulk Solids magazine.

Finding Information on the Internet

www.processingmagazine.com
This is the webpage for Processing magazine.

www.publist.com
This data base provides publication information for over 150,000 magazines, journals, newsletters and other periodicals.

www.zdnet.com/searchiq/subjects/
Directory of specialized search engines owned by company that produces technological magazines and newsletters.

Business/Competitor/Financial Information

www.about.com
About finds experts and expert advice on a wide range of topics including management, marketing, agriculture, composites/plastics, and many industries.

www.accessabc.com/
This site is a quick online service for the latest top-line circulation information from the ABC. It lets you sort and search circulation data by the publisher title.

www.activemedia-guide.com
U.S. Business Reporter is a private company that sells business reports and industry overviews.

http://adage.com/
This site allows you to search a database of 400 advertisers, and the agencies and agency networks that handle the accounts.

www.ajb.com
America's Job Bank is the biggest and busiest job market on the Internet.

www.alacra.com/alacra/index.htm
Alacra provides access to market reports from several sources.

www.annualreportservice.com
www.reportgallery.com
www.carol.co.uk
The above sites allow you to obtain annual reports for many companies.

www.anywho.com
Site provides residential and business phone numbers. It also allows you to reverse search using phone numbers instead of names.

Finding Information on the Internet

www.arcweb.com

ARC creates marketing reports which they offer for sale on various industries and products as well as customized reports; they offer a free newsletter which keeps you up to date on events in various industries.

www.bizjournals.com

This site is a major publisher of local/regional business journals whose articles can be found on-line.

www.bnet.com

Bnet offers a wide range of documents to help businesses solve problems with special focus on providing white papers.

www.brint.com/km/

Business and technology web portal (Business Research in Information and Technology); contains search engine and papers on various areas of business management.

www.businesscreditusa.com

This site provides a free basic directory and credit information on 12 million U.S. companies. Added information is available for a fee.

www.ceoexpress.com

This site is a cluttered but rich site having a strong emphasis on business news sites.

www.cordis.lu/en/home.html

This site provides information on Research and Development sponsored and supported by the European Union.

www.corporateinformation.com

Data base providing company information.

www.datamonitor.com

Fee based services providing business info and profiles on thousands of companies and industries.

www.dialog.com

Fee based access to 280,000 U.S. company profiles.

www.dnb.com

You can find credit reports available for a fee from D&B; wide range of services and products; good website to monitor company information.

Finding Information on the Internet

www.ecnext.com/commercial/knowledgeceter.shtml
ECNext Knowledge Center has compiled 250,000 sources from 500 publishers in searchable collections of market intelligence.

www.economy.com
Site provides free access to over 1 million economic and financial data series.

www.edgaronline.com
Tool that extracts information filed by companies to the SEC

www.eiu.com
This site provides access to business and news information and information on business environments in different countries.

www.exporthotline.com
This site allows you to find information on U.S. imports and exports.
www.factiva.com
Fee based tool geared for business information.

www.financials.com
Find free information on over 10,000 public companies.

www.firstresearch.com
This site offers more than 140 industry profiles. The profiles cover recent developments, trends, opportunities and other information.

http://fisher.lib.Virginia.EDU/active_data/index.html
GSD Center at the University of Virginia provides access to several economic resources. See the Center's home page for a complete list.

www.flipdog.monster.com
FlipDog is an online job site.

www.galegroup.com
Site provides access to several data bases containing market information.

www.globalspec.com
www.thomasregister.com
Sites allow you to search for manufacturers via product queries; links to manufacturers home pages; some biographical info on manufacturers. GlobalSpec allows you to search for engineering related articles and papers.

www.greenbook.org
Searchable directory that helps you to select the best marketing research services for your needs.

Finding Information on the Internet

http://groups-beta.google.com
www.hotbot.com
http://internetdatabase.com/
www.reference.com
These sites help you to find links to industry discussion forums.

www.gov/main/www/srchtool.html
Site provides a search tool for finding information on government web sites.

www.hoovers.com/free
Free thumbnail profiles on companies; fee based information provides more details.

www.integrainfo.com
Fee based data base of benchmarking data on privately held companies. Click "free stuff" and get up to five free snapshot reports on over 900 industries.

www.knowledgeexpress.com
Fee based; search for key contacts, market information, patent information, and new product concepts.

http://knowledge.insead.edu/home.cfm
This site provides abstracts, papers and case studies on business research topics.

http://knowthis.com
KnowThis is a collection of the best and most substantive links on all major areas of marketing.

www.knowx.com
This site provides access to public records on a pay-per-page basis.

http://library.hbs.edu/industries/industry_guides/
Baker Library industrial guide points you to sources of current and historical information on various industries

http://library.sau.edu/bestinfo/online/onlindex.htm
St. Ambrose University site links to collections of statistics from various sources.

www.marketresearch.com
www.mindbranch.com
Find indices, table of contents, and excerpts from marketing research reports.

www.netadvantage.standardandpoors.com
S&P's industry surveys cover 52 major U.S. industries.

www.onesource.com
OneSource provides industry and market information from many sources.

Finding Information on the Internet

© 2005 Monitor Technologies, LLC.

.
www.prnewswire.com
www.prnewswire.com/cnoc/cnoc.html
www.internetwire.com/iwire/home
http://home.businesswire.com
http://news.thomas.net
www.usnewswire.com
All the above are sites from which you can retrieve press releases from companies based on subject. Some require you to register but all are free.

http://research.thomsonib.com
For a fee Thomson offers industry specific statistics, economic indicators, trends and forecasts from 202 trade associations, covering many major industries. Thomson is also a good source for company information and market intelligence.

www.rh.edu/library/industry/industry.htm
Cole Library industry data finder provides likes for researching many industries

www.salary.com
Searchable data base by occupation and geographic area.

www.sba.gov
This site provides statistics about recently approved loans to small U.S. businesses.

www.sec.gov
Good starting place for information about specific companies and industry trends. SEC requires all publicly held companies and some large private companies to report all their financial information.

www.sedar.com
Canadian equivalent of Edgar system; database provides access to public financial information on Canadian businesses

www.sic-naics-codes.com
www.census.gov/epcd/www/naics.html
www.osha.gov/pls/imis/sicsearch.html
These sites allow you to find SIC codes and NAICS codes by industry

www.sixfigurejobs.com
www.monster.com
www.careerbuilder.com
Sites for executives/professionals seeking high paying jobs.

http://strategis.ic.gc.ca
Canadian website that provides industry information

Finding Information on the Internet
© 2005 Monitor Technologies, LLC.

http://tech-net.sba.gov
This is an electronic gateway of technology information and resources for and about small high-tech businesses. It is a search engine for researchers and scientists, a marketing tool for small firms, and links to investment opportunities and sources of capital.

www.10Kwizard.com
Interface to SEC's EDGAR database. It provides free access and search to the EDGAR system. You can view current and historical filings, from the start date of each company's existence by keywords, phrases, and names.

www.the-dma.org
Site offers information and resources on direct mail, telemarketing, lists, catalogs, interactive marketing, B2B, and international marketing issues.

www.topica.com
Topica contains a comprehensive searchable directory of more than 100,000 mailing lists.

www.tornado-insider.com/radar/
This site provides research start-ups, investors, and advisors to high-tech Europe. You can search for company press releases and profiles.

http://www.tripod.lycos.com/smallbiz/index.html
Fee based and some free company information.

www.tscentral.com
www.tsnn.com
www.conferenza.com
www.techweb.com/calender
www.expocentral.com/industrial/industry_industrial_supplies.html
Trade show central is searchable by industry, type, keywords, and geography; covers more than 30,000 tradeshows, conferences and seminars world wide. The other sites have the same functions.

www.valuationresources.com/IndustryReport.htm
Provides industry reports for about 200 industries by SIC code.

www.whowhere.com
Offers links to company home pages; allows search for company by SIC code.

www.worldatonce.com
Directory information on 40,000,000 firms worldwide. User can search by region as well as by SIC code.

Finding Information on the Internet

www.worldpages.com
Offers many information tools including SIC codes.

www.zacks.com
Search for information on companies whose stock is traded.

www.zapdata.com
ZapData provides free company profiles on 12,000,000 businesses and reports on many industries.

Keeping up Tools

www.copernic.com
Free and fee based on-line searching tool using multiple engines; organizes reports, etc.; can be used to do periodic scans on a given query and email results back.

www.factiva.com/infopro
Site that helps companies integrate news and business information into their daily workflow; download articles (i.e. "Value of free web-based information").

www.keepmedia.com
Fee based searchable data base for magazine and print articles. You can set up key terms and be notified by email of any hits. You can search the data base for free and see what is available.

www.getthepatent.com
Fee based site will perform automatic periodic searches of queries and email results.

www.individual.com
Provides news and information on selected industries; for a fee you can have specific targets watched automatically.

www.intelliseek.com
Fee based; tap 300 search engines and 600 databases to find, analyze, filter, report and track information.

www.lexisnexis.com
Fee based site that allows you to plug in names of companies you are tracking and have articles faxed to you as they come up in the database.

www.patentalert.com
This site keeps you current with new patents filed in your field of interest via email when the patent is awarded.

Finding Information on the Internet

www.smartbrief.com/index.html
Industry news alert service sent to your email. Obtain headlines and abstracts of recent articles that relate to the industry selected.

www.thomasnet.com
You can register at this site and establish key words. ThomasNet will then send you emails with articles that reflect the key words you have established.

www.tracerlock.com
Free and fee based monitoring of search engines, newsgroups, etc.; notifies you when it finds matches for you search terms

Competitive Intelligence/Researchers
www.aiip.org/memberdirectoryhome.asp
www.burwellinc.com/bedirectory.html
These websites are useful for locating for-hire researchers including CI people.

www.cognis.com
Site that offers fee based software for gathering information from competitors and analyzing it.

www.competia.com
Site dedicated to CI.

www.factiva.com/cicenter
Links to recent articles o CI; also article archive and email alerting service

www.fuld.com
Website of CI consultant; contains downloadable documents & articles; services include literature research, collection management, consulting, training, new monitoring, etc.

www.fuld.com/Tindex/I3.html
Find CI on a business competitor or company for purchase. Directory provides links to over 600 sites that CI researchers should be aware of.

www.scip.org
www.scip.org/news/cimagazine.asp
Website for the Society of Competitive Intelligence Professionals; second site is an online magazine dedicated to CI information.

Patents/Trademark Tools
www.delphion.com
Delphion Intellectual Property network provides free access to a wide variety of data collections and patent information.

Finding Information on the Internet
© 2005 Monitor Technologies, LLC.

www.derwent.com
www.derwent.co.uk
Searchable patent and trademark data bases (fee based).

www.e-gineer.com
Find domain names, trademarks, definitions, synonyms, etc.

www.getthepatent.com
Fee based site with guest try-out offer for 5 patent searches.

www.knowledgeexpress.com
Fee based; search for key contacts, market information, patent information, and new product concepts.

http://lib.ucsd.edu/se/list.html?type=3
This site offers a patent searching tool.

www.patentalert.com
PatentAlert keeps you current with new patents filed in your field of interest via email when the patent is awarded.

www.patentcafe.com
This site allows you to find patent information.

www.questel.orbit.com
Search for patent and trademark info in U.S. and internationally.

www.thomson-thomson.com
Fee based service allows you to search for info on trademarks and copyrights.

www.uspto.gov
www.uspto.gov/patft/
Website for US patent office; offers a searchable data base.

Professional/Industry/Organization Locators
www.about.com
About finds experts and expert advice on a wide range of topics including management, marketing, agriculture, composites/plastics, and many industries.

www.ansi.org
Search engine to locate standards from premier U.S. standards organization.

Finding Information on the Internet

www.asaenet.org/
www.associationcentral.com
www.ipl.org/ref/aon
These sites offer links to professional and industry organization websites.

www.asme.org
The website for the American Society of Mechanical Engineers.

www.expertisesearch.com
This site is a national and international guide to experts/consultants in all areas.

www.galegroup.com
Directory of trade and professional organizations (fee based).

www.isa.org/
The website for the Instrument Society of America.

www.iso.org/iso/en/isoonline.frontpage
Search engine to locate standards from premier international standards organization.

www.specialissues.com/lol
Data base of ranked listings of companies, people and resources available for free on the Internet.

Marketing Report/Tools
www.arcweb.com
ARC creates marketing reports which they offer for sale on various industries and products as well as customized reports; they offer a free newsletter which keeps you up to date on events in various industries.

www.ecnext.com/commercial/index.shtml
Use this site to search for market reports. You must purchase any reports of interest.

www.frost.com
On-line access (for a fee) to marketing reports from Frost & Sullivan.

www.gartner.com
Gartner is a consulting firm that produces market research reports; summaries and briefing papers available for download.

www.greenbook.com
This site is a free directory of market research firms; good for identifying companies in specific industries.

www.marketresearch.com
Contains aggregate of market research reports from various publishers; reports offered for fee; no charge to search data base.

www.marketsearch-dir.com
Free information on 20,000 market research reports from over 700 firms; offers a summary for free and information on where to get full report; searchable by industry.

www.mindbranch.com
This site offers market research reports (for a fee) from various sources.

www.technical-insights.frost.com
Site offers market research services from Frost & Sullivan and publishes several newsletters on new technologies, R&D, new patents, technology impact analysis, etc.

Law/Legal Tools
www.findlaw.com
Search engine to find information on legal issues.

http://lawcrawler.com/findlaw.com
LawCrawler finds information on litigation in the US.

www.uscourts.gov/
This site allows a query of all publicly available federal judicial documents.

Other Tools
www.amazon.com
Amazon is more than just a book seller. If you enter key words using its search engine you can get leads to some good websites. Amazon also has market research reports for sale (which are resold from third parties). In addition, Amazon can provide full search text from tens of thousands of books for free.

www.amusingfacts.com
This is a good site to find trivia type of information.

www.bartleby.com/quotations
Columbia University site that contains collections of quotations.

www.benedict.com
Good site for addressing issues involved in the copyright of digital materials.

www.businesscreditusa.com
This site provides basic background info on 1,200,000 U.S. businesses.

Finding Information on the Internet
© 2005 Monitor Technologies, LLC.

www.convertit.com/go/convertit/measurement/converter.asp
www.onlineconversion.com
These sites have easy-to-use measurement conversion calculators.

www.cybergeography.com/atlas/atlas.html
www.mapquest.com
All the above sites offer various kinds of maps.

www.faqs.org/faqs
FAQS is a site that contains indices of Frequently Asked Questions on many subjects.

www.icann.org
Internet Corporation for Assigned Names and Numbers coordinates domain names.

www.invisible-web.com
This is a website created by the author of a book of the same name. It is a good site to find valuable information on how to find information on the Invisible web.

www.ietf.org
Internet Engineering Task Force establishes internet protocols and irons out problems.

www.isoc.org
This is the website of the Internet society that acts as a clearinghouse for technical standards.
www.ithaca.edu/library/training/icyousee.html
Thorough list of things you can do on the internet compiled by Ithaca College library.

http://iws.ohiolink.edu/companies
Obtain how to guides for company and industry research

www.lib.berkeley.edu/teachinglib/guides/internet/findinfo.html
This site offers a series of tutorials on all aspects of using the internet.

www.lib.purdue.edu/vlibrary
This is the webpage for Purdue University. It has dictionaries, phone books, science information, maps, and other reference materials.

http://marylaine.com/neatnew.html
This site allows you to find the latest new sites for research on the web.

http://owl.english.purdue.edu/handouts/print/research/r_quotprsum.html
This is a site that gives tips on professional writing; also tutorials on searching the internet, search engines and other topics.

Finding Information on the Internet

www.specialissues.com/lol
Data base of ranked listings of companies, people and resources available for free on the Internet.

www.techweb.com/encylopedia
This site provides definitions for 20,000 informational technology terms.

www3.best-in-class.com/database
Follows best practices for business in several fields like e-business, knowledge management, sales and marketing. Free searching for fee based documents.

www.unc.edu/cit/guides/
Contains guides on finding information on the web on education and technology.

www.w3c.org
Web consortium acts as a forum to discuss specifications, guidelines, tools and software.

www.websearchguide.ca/netblog
This site contains Gwen Harris's weblog on internet usage and searching.

www.zdnet.com/
www.cnet.com
www.jumbo.com
www.shareware.com
The above sites offer extensive software libraries – both free and for a fee – that you can download.

References

1. "Google Pocket Guide"; Calishain, Tara, Dornfest, Rael, Adam, D.J.; O'Reilly and Associates; 2003.
2. "Find It Online"; Schlein, Alan M.; Facts on Demand Press; 2003.
3. "The Skeptical Business Searcher"; Burkman, Robert; Information Today Inc.; 2004.
4. "Extreme Searcher's Internet Handbook"; Price, Gary and Hoch, Randolf; Cyber Age Books; 2004.
5. "Web Search Garage"; Callahan, Tara; Prentiss Hall; 2005.
6. "The Invisible Web"; Sherman, Chris; Cyber Age Books; 2001.
7. "Introduction to Online Marketing and Industry Research"; Shamel, Cynthia; Thomson Learning; 2004.
8. "Introduction to Online Company Research"; Dobson, Chris; Thomson Learning; 2004.
9. "Web Based Analysis for Competitive Intelligence"; Vibert, Conor; Quorum Books; 2000.
10. "Super Searchers on Competitive Intelligence"; Carr, Margaret Metcalf; Cyber Age Books; 2003.
11. "Early Warning"; Gilad, Ben; American Management Association; 2003.
12. "Competitive Intelligence: A Guide for Your Journey to Best Practice Processes"; Leavitt, Paige and Hasanali, Farida; APOC; 2004.