III.

Interesting Web Resources: http://www.tveyes.com or http://ehow.com or http://iltrafficalert.com (compare with http://traffic.com) or http://kiva.org or http://emergingmed.com

Looking for Information Online

There is a big difference between looking for information online and information in printed published form. Printed materials enable us to easily examine the following:

- Format: the form in which info is presented (magazine, paperback, hardcover) and the level: (scholarly, popular, juvenile, etc.)
- Scope: how clearly, cohesively and extensively the info is presented.
- Relation: to other works in the same field, reflecting how it builds upon the discipline's body of knowledge.
- Authority: the credentials of the author and the reliability or veracity of the source of the info.
- Cost: does the value justify the price

Online publishing differs when you consider:

- Format: design and layout of the information can easily overwhelm or disguise content.
- Scope and Relation: Comprehensiveness is difficult to determine. Individual pages rely on links to provide scope. Online doesn't mean up-to-date or current.
- Authority: user must look carefully for statements of responsibility
- Cost: not a big factor in getting online material published

The Importance of the Critical Thinking in Relating to the Internet

The dictionary says information is the communication or reception of knowledge or intelligence. It is the process through which an object or knowledge is impressed upon our minds that brings about a state of knowing. You are thinking critically when you are actively thinking about your thinking. Through this reception and impression - and self-examination - information becomes useful knowledge - sounds a bit like learning, doesn't it?

Where does information come from? From data being transferred. Data: something that is given from experientially encountered or from being admitted or assumed for a specific purpose - comes from scientists, researchers and other observers - We can get our info from our own world of sources:

We can conceive of five 'rings' of info surrounding us: Central is the 1) internal info of our reactions to our environment. Surrounding that is 2) 'conversational' info - which we pick up around us - also through email or online chat. 3) 'Reference' info is the next ring - the materials where we can seek out the systems of our world from a cookbook to online resources. 4th ring is News - transferred to us via the media. Final ring is Cultural Info - history, philosophy, the arts and our attempts to understand and put meaning into our civilization.

More info does not necessarily lead to a better or more satisfying life. It can lead to 'gluttony' or anxiety. Information consumes our attention. A wealth of info creates a poverty of attention and a need to allocate that attention efficiently among the abundance

In this class, we will chart a pathway through online information and introduce tools and concepts to help us gain value from our online experiences.

Evaluating Web Sites

As online surfers, we need to keep three elements in mind when examining information:

1) **examine the evidence** (information) presented;
2) **look for credible sources**, and
3) **evaluate the assumptions** and logic in the material presented.

Remember, there is no single authority on the Web nor any regulation of the information provided. Anyone can create a web page and make it available. Therefore, you should view the information you find on the Web with a certain amount of skepticism and a critical mind. Many of the assumptions about information delivery are no longer valid because the source and credibility of information are much harder to discern online.

We need to differentiate between fact and opinion. Examine assumptions, including our own. Be flexible and open-minded as we look for explanations, causes, and solutions to problems. Stay focused on the whole picture while examining the specifics: be aware of fallacious arguments, ambiguity and manipulative reasoning.

We can look at eight items to determine the validity of this information:
Source (who is the author), Credentials (of the author), Type of Information (Scholarly, Popular, etc.), Purpose, Timeliness (current), Style, and Assumptions (of the author.)

A lot of Web pages look very good on the surface but may not contain reliable information.

Basic criteria for evaluating Web sites

Who **sponsors**/ supports the Page/Site:

Is it a commercial vendor trying to sell you a product?
From an educational site?
Is the page really supported by the institution or just a page put up by a student?
Consider sites from government agencies and professional associations.

When was site created and/or updated?

Is information **current**?

A well-constructed site should contain a creation/update date.
Is the information updated on a regular basis or just created and never maintained?

What are the Credentials of the Author?

Is the author, contact person, or webmaster identified?

Searching for Right Search Tool

Researchers now have it all on the web: facts on virtually anything, available from anywhere, unfiltered by reporters, editors or publishers and most of it free. But sometimes we feel we have too much info - often way too much and it may not be correct.

Imagine you're at a library - you're interested in vintage Ford Mustang convertibles, so you'd like to find a book with lots of pictures about the car. You spot a librarian and ask him for some help. Bad news: this person understands only Russian, which is not your native language, so you reduce your query to some basic words: Book, Car, Ford. The librarian seems to understand and trots off to fetch a book - he

brings back many volumes but not satisfy - some are on cars but not exclusively Mustangs, some are about horses, some are about former President Ford. You spot another librarian - this one speaks only Chinese - and you begin again!

This frustrating scenario is what it's like to search for info on the web today. The web contains some 1.2 billion pages and roughly doubles in size each year - so is enormous and full of non-relevant information to your needs. 7 of 10 users are dissatisfied with search engines - many are not so good but most don't know how best to use them (less than six percent know how to use Boolean search terms - "and" "or" plus & minus signs. Research found most frequently occurring search inquiry is fragment of a URL, which indicates most can't tell their browser from a search engine.

A survey by WebTop shows most Americans experience 'search rage' if they can't find what they're searching for within 12 minutes!

Why Search Stinks.
Most search engine functions have been dumbed down to produce lots of results for unskilled searchers: a common problem. - leaves user unable to find what wanted and wondering how to narrow down the search field. - relevancy is key.

Two basic approaches - using man or machines. Machine approach holds that technology can crunch great numbers of pages that no human could process to find what you need. AltaVista, Inktomi, and Excite were the original heavy hitters. But despite hype, search engines still can't distinguish among web pages based on their content. 'Man' approach uses best supercomputer ever invented: brain. No tech. exists that can beat ability of person to select highest quality web pages and discard the rest - see Yahoo. The only way you can pinpoint info is to learn how to do efficient searches and which tools are the best for which purposes (you can choose from hierarchical indexes, standard search engines, alternative search engines, meta search engines, and databases)

Search tools are wildly competitive, with nearly a hundred solutions out there.

A Survey of Search Technologies

Boolean: Retrieves documents based on the number of times the keywords appear in the text. Simple expressions like And, Or, and Not make queries more specific. Virtually all search engines combine Boolean procedures with other methods.

Clustering: Dynamically creates 'clusters' of documents grouped by similarity, usually based on a statistical analysis of the contents and the structure of the document text. Vendors include www.vivisimo.com.

Linguistic analysis: Dissects words using grammatical rules and stats. Finds roots of words, alternate tenses, equivalent terms, and likely misspellings. Also called stemming, morphology, synonym-handling, spell-checking. Virtually all tools use all or some of this.

Natural language processing: Uses grammatical rules to find and understand words in a particular category, like product names. More advanced approaches classify words by parts of speech in an attempt to understand their meaning. Also called named entity extraction, semantic analysis. Vendors: www.albert.com, www.inxight.com, www.inquira.com

Ontology: Formally describes the terms, concepts, and interrelationships in a particular subject area. A vocabulary needed by systems in order to associate and connect data across multiple databases. Also called knowledge representation. Vendors: www.endeca.com, InQuira, iPhrase.com, verity.com.)

Probabilistic: Calculates the likelihood that the terms in a given document refer to the same concept as the terms in the query. Also called belief networks, inference networks, Naïve Bayes. Vendors: Autonomy, www.recommind.com, Microsoft.

Taxonomy: Establishes the hierarchical relationships between the concepts and the terms in a particular search area. Also known as categorization. Vendors: Gammasite, H5technologies.com, Yellowbrix.com.

Vector-Based: Represents documents and queries as arrows on a multidimensional graph and determines relevance based on their physical proximity in that space. Also known as vector support machine. Vendors: convera.com, Google.com, Verity.

Unstructured Information Management Architecture: help other programs acquire and analyze text, audio and video and arrange them in structured forms. IBM

Internet Search Topics
What are people searching for? In rough order, the top terms are: music, travel, sex, games, and eBay

Types of Internet Search Tools
There are three main types of search sites: various combinations of these keep appearing. Each search tool takes your request (the search items you type in) and retrieves a set of records that match.

Type 1: Subject Directories (Also known as Hierarchical indexes): Lists of Internet resources, arranged by subject.
In a Subject Directory: Yahoo.com was one of the originals - people trained to categorize information, such as librarians and indexers, examine web sites and put them in categories and subcategories. Thus, when you do a search here, you are more likely to find what is relevant to what you are looking for… Librarians maintain a directory at http://lii.org with more than 16,000 sites. Drawback: they are selective, and because created by humans, they can include only a tiny portion of what's out there. Yahoo uses a 'standard search engine' as well. So a search is split into several sections: "Category matches" tell you if your topic matches one of Yahoo's existing categories. " Site matches" are those that have been indexed and categorized. "Web pages" provide links to pages located by the search engine. Also "related news" and "net events" Use this type when you are looking for **very broad or general topics**.

Subject Directory Summary
- Subject directories are specialized Web search tools that select other web sites and organize them under broad subject headings.
- These search guides are compiled and maintained *manually.* People can also build the catalog by registering their pages with the directory.
- Due to the Web's immense size and constant transformation, keeping up with the important sites in all subject areas is impossible. A guide compiled by a subject specialist to resources in his area of expertise is valuable for locating relevant info.
- No two subject directories categorize subjects in the exact same way.

- When you are looking for general info or browsing a board topic, subject directories are an excellent place to start.
- You will generally need to use several subject directories to be comprehensive.
- Subject directories cover only a small fraction of the total Web pages available.
- Yahoo, the largest and most popular subject directory covers less than 3% of the entire web.

- Major subject directories: many library home pages.

Type 2: Standard Search Engines: The Comprehensive Indexes of the Internet

Search engines use computer power to build their databases. Search engines employ software programs called spiders or robots which automatically 'crawl' from web site to web site continually copying content. New or changed info is send back to the search engine. The retrieved keywords are built into searchable indexes. The engines then calculate mathematically how relevant the pages are to your search terms - using its own algorithm to rank pages. Factors in the calculation include frequency and placement of keywords on a page and their occurrence in the description that owners write of their pages (which are invisible to users). The engine puts the pages with the highest score at the top of their list. Each item in the index points exactly to where the web resource is located on the web. Use a search engine to find a **concept or a specific phrase**. Avoid these when you have a very broad topic.

While all search engines are intended to perform the same task, each goes about it in a different way. The search algorithms are kept secret. Some record all pages; others are selective. Some record only titles; others all text at the site. Search engines differ in speed, design, the way they display results, the help they offer. To use a search engine, enter a term or terms; search engine returns a list of documents containing the term(s).

Most search engines display a description of each site when listing the search results (often the first 200 words)

Some search engines are very comprehensive, some are very selective, prioritizing what they index by popularity or frequency of updating of how often a term appears in a document.

In an attempt to get the most relevant results at the top of the list of items you retrieve, search engines apply 'ranking Algorithms' which seek to display the most relevant items first.

Plan B: if you enter the same search on several search engines, the results will vary widely. Get into the habit of using several Internet search engines. Knowing alternative search tools will also be very useful as a backup during "internet rush hour."

Major search engines include Google and Bing; There are lots of former strong-holds: AltaVista, (owned by Yahoo, formerly paid ad provider Overture, and before that Digital Equipment Company (DEC)) Excite, HotBot, Inktomi, Ask, etc. They generates much revenue by 'sponsored search' that serves up connected advertising with search results. Yahoo has alternatively owned and outsourced its web search so not to have to share revenue with a third party like it previously did with Google and Microsoft.

Lycos.com was one of the first search engines, launched in 1994 at Carnegie Mellon U. AltaVista should have been Google. It came along in December 1995 and revolutionized search. Back in the day, web crawlers visited URLs, captured headers and headlines for sorting in their indexes. Crawlers had to wait

several seconds to query a site. At best, that mean an engine could cover tens of thousands of sites a day without picking up changes. AltaVista had a multithreaded crawler named Scooter which could ping, track and respond thousands of sites as separate threads. It revealed the size of the Web back then at 16 million pages – two months later, it was 25 million. Today there are billions of active sites – each with dozens to thousands of pages. Digital (and later Compaq, later HP) never knew what to do with AV – spun it off, then bought it back, sold it again (finally to Yahoo).

Alternative search engines take various approaches to ranking & sorting and are often more helpful than standard engines. For more information and listings, see http://searchenginewatch.com

See http://mashable.com/2013/02/26/alternative-search-engines/

Ask (www.ask.com) has constantly sought to relaunch itself back to relevance. Originally known as Ask Jeeves, Ask allowed you to type a question in plain English - it recorded millions of questions and has found web sites that answer those questions. Ask pioneered giving you the answer rather than a link to the answer. Ask was one of the first to link all its sub-brands to offer **local search**.

After getting your question, Ask scans its database of questions and answers - it then gives you a list of questions it 'thinks' you want answered. If you select one - it lists sites with the purported answers. Ask ranks sites based on the number of topically related pages that link to it. Recent Ask upgrades give little thumbnails of actual websites instead of requiring you to click through, allows users to store individual search results, organize them and share them, lets users expand or narrow keyword searches based on conceptually-related topics (Zoom) and improve its question/answer technology. This group was owned by Barry Diller – owner of match.com, evite.com, citysearch.com, Hotels.com and recently-spun-off Expedia.com. The old 'Jeeves' symbol is long gone away.

Google takes another tack - it first matches your keywords to pages its collected in its index. Then, however, it ranks each page based on how many other pages link to it - and how many link to those pages in turn. Essentially, Google assigns a numeric score (up to 10) to every page it crawls. The score is based on how many other sites link to that page as well as the Page Ranks of those sites and it influences how high the page appears among a set of search results. The pages you see at the top of results are those with the highest number of links to other pages. The idea is such popularity is meaningful - just as a diner with lots of trucks parked in front of it serves better food than another with an empty lot. Google stores the data it indexes in special servers which can access its data much faster than other engines. See http://googleguide.com

Among the company's newest services is one that draws upon newly developed algorithms to list the academic research that appears to be most relevant to a search request. Google doesn't plan to charge for the service nor use the feature to deliver text-based ads, the primary source of its profits. Although Google already had been indexing the reams of academic research online, the company hadn't been able to separate the scholarly content from commercial Web sites. By focusing on the citations contained in academic papers, Google also engineered its system to provide a list of potentially helpful material available at libraries and other offline sources. The scholarly search effort continues Google's aims to probe even deeper into content available online and offline. Last month, Google expanded a

controversial program that invites publishers to scan their books into the search engine's index, enabling people to peek at the contents online before deciding whether to buy a copy.

Serial entrepreneur Bill Gross, developed the search advertising model that Google eventually copied and now relies on for most of its profits (Google eventually paid a licensing fee to Overture Services Inc., the company that Gross created and is now owned by Yahoo Inc.)

See http://www.gizmopod.com/gallery/6196/15-amazing-hidden-features-of-google

See Guide to Google Drive: http://www.educatorstechnology.com/2014/01/the-comprehensive-google-drive-guide.html

As the Internet was just beginning to grow in the mid-90s, two Ph.D candidates in computer science at Stanford began exploring the relationships between pages on the Internet. What started as an unassuming research project yielded a popular system for ranking web pages. By Jan. 1996, the pair, Larry Page and Sergey Brin had begun collaboration on a search engine called BackRub, named for its capability of analyzing the back links pointing to a web page. Google first emerged when Page and Brin bought a terabyte of disks and built Google's first data center in Page's dorm room. Page and Brin used Duplo blocks as the ccasing for 4GB hard disks for their first server. (Google is a play on 'googol' a math term for a 1 followed by 100 zeros – a reference to organizing the seemingly infinite Web). Google is fueled by more than 20 innovations in search technology: the most notable is the PageRank concept of analyzing a page's link structure to determine its value. Links coming from pages considered more important or authoritative on a given subject are given more weight and help increase the rankings of other pages. This is complicated by the multiple languages, encodings and data types of the world's pages. Google answers 150 million queries a day and its collection of web documents, usenet discussion posts, images, catalogs, PDF files and Word documents is more than 5 billion - the largest of its kind.

By 2002, Google handled 14% of all searches worldwide, followed by MSN's 13% and AOL's 5%. Today, Google is the most popular place to find things on the Internet. Looking at it another way, last year, Google handled more searches than most of its competitors combined. Though China is a different story.

Yahoo played a vital role in Google's rise by originally licensing Google's software to run the searches on its portal. Most of Yahoo's original results were generated by Google's software. Today, Google handles about 69% of all search-engine-generated traffic to Web sites. Yahoo stopped trying to supplant Google with a rebuilt search engine running off acquisition Inktomi's technology; instead it partnered with Microsoft's Bing. Yahoo wants some of Google's advertising money. Google still makes most of its money through advertising – it gets paid to put small text ads above and next to non-commercial search results. Its 200,000 advertisers make it the world's largest Web-ad network.

Microsoft's percentage of the search market, with its Bing, has grown largely at Google's expense.

See http://mashable.com/2013/07/05/google-tools/

Google Hacks For a list of all Google specialty sites, see http://www.google.com/help/features.html

- Google is not case sensitive and it limits searches to 10 words – ignoring anything beyond.
- Do you have Google Voice - offering one phone number "for life," with free calls across the U.S., and voicemails transcribed and sent as e-mail or text messages.
- See Search Console at http://www.google.com/webmasters/sitemaps
- If you are unable to find a webpage that was on a site a few days ago but has vanished, try the prefix: **cache:**
- http://translate.google.com
- Type in the ticker symbol for any stock and get a real-time quote and related business information.
- Type in the routing number for your FedEx or UPS package and track its progress.
- Other Google Number searches allow lookups for a car VIN number and aircraft tail number, FCC numbers printed on gadgets and patents: Try this search "patent 5123123"
- Type any arithmetic, trigonometry, weights and measures, or conversion problem and get the answer .
- Type in airport code, airline and flight number (ex. ORD United 241) to get arrival and departure.
- Type in airport code to get local weather and delay info or type in the word 'weather' and a city name.
- Type 'movie:' with the name of the movie and find a compilation of reviews.
- Type in a dollar / currency amount like $500 USD in euros nd Google will give you the conversion
- See http://google.com/calendar
- Type in a phone number: 10 digits, no prefixes, no spaces, and find the name and address – and click straight to maps and driving directions. Click on the telephone icon to the left of your name and you can submit a form so that Google removes your name from their directory.

The result page for phonebook: lookups lists only five results, residential and business combined. The more specific rphonebook: and bphonebook: searches provide up to 30 results per page. For more chance of finding what you're looking for, use the appropriate targetted lookup.

Google's News Search supports two special syntaxes.

intitle: Finds words in an article headline. intitle:miners

site: Finds articles from a particular source. Unfortunately, Google News does not offer a list of its over 4,000 sources so you have to guess a little when you're looking around. miners site:bbc.co.uk

- To limit searches to a particular site, try Site: with the URL suffix such as site:gov statistics
- Use –site: to exclude a type of site.
- To search for a specific type of file, such as Adobe PDF or Microsoft Word, include 'filetype' with the suffix such as filetype:pdf global warning
- To search for pages that may have been removed from a website, include 'cache' with the domain, such as cache:www.house.gov bill

Google's Library Project

Google is partnering with libraries at Harvard, Stanford, Oxford, University of Michigan and the New York Public Library to digitize some of their collections. More than 50 million books are part of the

project, expected to take up to a decade to scan. Indexing library books will make Google more competitive. The first wave of books are in the public domain and authors with out-of-print books can sign up their works at http://print.google.com Searching Google will also yield links to buy the book in question. This project is growing in controversy as Google forces publishers to opt-out books they wish to disclude

Google's Future

See https://www.thinkwithgoogle.com/platforms/display.html
Google faces an uphill battle to stay atop the search engine heap. But keeps branching out with Android, Google Doc, Google Drive, and its Chrome browser and operating system.

Pay for Placement
In the old days, say about 5 years ago, if you searched on a keyword, you would find the most relevant matches. Now, companies are playing both for position and even inclusion in search results. This means the search engine company will at least guarantee that you will be in its database. This is similar to what the Yellow Pages has been doing in print for years. On most search engines, the first 3-10 hits will be sponsored links, whether they are identified or not. The FTC strongly suggests sponsored links be labeled as such. If this continues, you can only imagine how bad it might get…but there's no fighting it.

$1.90 is the average keyword price paid by online advertisers – a decline of 4%

History of Search

Phase I
In this era, during the late 90s, Altvista, Excite, Hotbot, Infoseek, Lycos, WebCrawler debuted, allowing simple text-search to find and rank pages. Searching for Nerf Ball would find lots of pages with Nerf and Ball. Marketers decoded these algorithms to game the system. That led to…

Phase II
The Age of Google. Moving beyond text search, Google's PageRank system ranked pages according to a series of sophisticated off-the-page criteria, including the number and quality of links to a given page. There are limits to this one-size-fits-all system. Leading to a nascent…

Phase III.
Blends vertical, real-time or specialized search results via mobile devices.

What's Hot?

See http://searchenginewatch.com

http://www.wolframalpha.com/ - not quite a search engine. More of a service – completed answers to queries using enormous collections of data it has amassed. Includes three course assistant apps to support Algebra, Calculus and Music Theory. See http://www.businessinsider.com/awesome-things-you-can-do-with-wolfram-alpha-2013-7

http://bing.com – Microsoft's latest anti-Google effort. See Best Match and Deep Links. Enables video searching

http://mahalo.com – employs human editors – wiki-like community-generated content review of results.

http://chacha.com includes an option to direct queries to live search guides who get paid based on how many questions they answer using chat and software that direct queries to people with domain-specific experience. Guides are supposed to answer in 20 seconds while the user sees video ads. ChaCha addresses a huge problem of credibility and authority in social search applications like wikipedia

Amazon.com's search tool at http://www.a9.com basically consists of Google's database and algorithms along with some nice features like access to Amazon's titles, the Internet Movie Database, Answers.com's reference information. Google also provides the bulk of A9 search results, ads and revenue through an ad-sharing agreement. On the right side of the screen, you can create a personal history of your searches – and bookmarks – that are accessible from any web browser. This personalization is the latest trend in search tools. A9 also has a feature called OpenSearch. Check it out.

Type 3: Meta-Search Tools

Meta-search tools are web sites with many different search tools on one page which allow you to conduct a search without having to go to each search tool location individually. They are essentially search engines that search other search engines These sites contain many different search tools on one page, but you must search each tool *separately.* Use a meta-engine for a very obscure or '**needle in a haystack'** topic.

Meta-search engines:

- Search multiple search tools simultaneously.
- 'Borrow' the indexes of other search engines.
- Are convenient but allow less control of your search than individual search engines. You do not have all the advanced searching features that can be found with individual engines.
- Generally display the top ten answers from each of the individual search engines.
- Number of search engines covered varies.
- Some eliminate duplicates.
- Some integrate results.

Example: info.com – which combines the top links of Google, Yahoo, Ask, Inktomi, Alta Vista and others. Concept behind meta-search: since no one single engine indexes the entire web, using a meta search engine can scan more sites. Downside: needs a 'lowest common denominator' search statement to search all of its engines. Can't handle complex searches.

In general, search engines do not touch databases – http://www.enth.com searches databases – mostly related to sports and demographics.

What do you think of visual representations of search results - on map-like screens that display results graphically: thebrain.com

Encyclopedias: http://www.encyclopedia.com
Http://wikipedia.com has become ubiquitous with its unique and controversial wiki-based ability for users to edit entries. In response to this sense of how does one assess experts, Larry Sanger, one of Wikipedias' cofounders is launching an alternative with vetted, non-anonymous contributors: http://citizendium.com

The Future

See the End of The Web, Computers and Search as we know it:
http://www.wired.com/opinion/2013/02/the-end-of-the-web-computers-and-search-as-we-know-it/

We're close to the day when search engines will offer enhanced service on an informed consent model similar to what Amazon.com uses for recommendations. By Christmas, we can expect a company that commands a large audience to offer personalized search as an option for users in exchange for registration and the willingness to have their search activities monitored and measured.

How much about our tastes and ourselves are we willing to furnish to search companies in the name of enhanced performance?

Find it Faster: Internet Search Strategies
"It's amazing that the amount of news that happens in the world every day always just exactly fits the newspaper." Jerry Seinfeld.

Using the Internet to search for something should be as simple as typing in a query and obtaining relevant results. In reality, Internet searching can be an exercise in frustration.

Searching with an Exact Address
- Remember every Web document has an address called a URL.
- The fastest way to retrieve a Web document is to type in its exact address (or to activate it via a Bookmark or from your Favorites list).
- You must type in the exact address; correct spelling and case is crucial. Many URLs are case sensitive.

The easiest way to find what you are looking for is to go to a dedicated site.

Bookmarks/Favorites
- When you find a useful Web site, save the address by setting a bookmark.
- Most browsers allow you to bookmark a site – to save a shortcut to the address. Click on Bookmarks or Favorites on your browser menu and choose Add Bookmark or Add Favorite.
- The URL and title of the currently loaded page will be added to your bookmark list.
- You can personalize your bookmarks by renaming them or by creating folders to organize your bookmarks by subject area.

  Some favorite Websites:
  http://www.buzzfeed.com/jessicaprobus/33-amazingly-useful-websites-you-never-knew-existed

http://mashable.com/2013/11/29/reddit-favorite-websites/

Searching with Web Search Tools

Because the Web has millions of pages with no central organization, individuals and organizations have developed a wide variety of search tools to assist users search for information. Although some of these tools are very comprehensive, there is no single search tool, site or method that will get you everything on the 'net.

General Tips

- **Analyze your info. need:**
  Is the Web the most appropriate place to start?
  Do you want to explore a subject or find an answer to a specific question?
  What organization is most likely to produce the information?

- **Where should I start:**
  Is the information you need related to a GENERAL TOPIC (company info, weather, web design, etc.) Choose a SUBJECT-BASED DIRECTORY like Yahoo

  Are you looking for a CONCEPT or a SPECIFIC PHRASE or NAME (e.g. maltese falcon, bulletproof vests) Use a SEARCH ENGINE.

  For a "NEEDLE IN A HAYSTACK" search, start with a META-SEARCH TOOL.

  For CURRENT information: consider searching newsgroups through groups.google.com

- **Try different search engines**. The same query will turn up completely different hits.

Power Searching

More and more search tools are adding special menus and aids to simplify advanced searching. These menus will guide you into using the principles in this section…

You *can* just enter a keyword and start looking at the results (200,000 results matter, you don't have to look at all of them).

- Retrieving a lot of hits does not guarantee that you will retrieve the best sources of information on your topic!
- A particular keyword phrase can bring excellent results on one search tool and poor results on another.

Picking your keyword(s) or search terms:

- Be very careful about SPELLING, PUNCTUATION, SPACING (DePaul vs. Depaul). Most search engines ignore case.
- HIGH-FREQUENCY TERMS (internet, computer) make poor search keywords.
- Do not use "SMALL WORDS" (in, the, a, to, of); extremely common words, STOPWORDS are ignored.
- Be aware of words that have MULTIPLE MEANINGS

Putting Your Keywords Together:

Default Search Logic: to better understand your search results. Most search engines default to "fuzzy logic."

Search Defaults

All searches use AND as the default linking operator between all of the search terms. Thus, searching for:

red herring

is the same as searching for:

red AND herring

For both of these searches, only those sites with "red" and "herring" in the site name or description will be returned. Sites that mention only "red" but not "herring" will not be displayed.

AND: all the words or word strings must appear. The Plus sign (+) is the same as AND

To get sites containing either "red" or "herring" use the keyword OR.

Boolean Searches

There are three main Boolean terms known as ''operators' to choose from: OR, AND, and NOT. Terms linked by the AND operator will return only those sites that match all the search terms linked by the AND operator. The default used by most search engines is AND. Use capitals for these operators as some sites require it and all recognize capitals. If you don't use any Boolean operators, only those sites that contain at least one occurrence of each search term will be returned.

Terms linked by the OR operator will return those sites that match any of the search terms linked by OR. For example,

grey OR gray AND parrot

OR: any one of the words or strings in the query could be in the result.

Terms linked by the NOT operator will exclude all sites that match the search term following the NOT. For example,

random andnot house

will find sites about randomness but exclude sites about the publisher Random House.

NOT: word or string must not appear. (The Minus sign (-) is the same as NOT.

Shorthand Search Terms

You can prefix search terms with "-" and "+" to force the exclusion or

inclusion of that term. This is really just shorthand for using the ANDNOT and AND Boolean operators.

Phrase Searching: Sometimes the order of the search terms matters. Using phrase searching can greatly reduce the number of sites that are matched by a search. For example, if you searched for

"Tour de France"

You would get only sites containing all three words "tour", "de", and "France" - in that order.

When searching a phrase or search terms that are to be taken together as a unit, enclose the phrase in quotation marks: *+"banana bread" (+nonfat or +"low fat") + recipe**

Treasure Hunt
Choose any search tool.
Answer the following questions:
Find out the top three American manufacturers of steam locomotives.
Find a good introduction to cross-stitching.
Find the population of Jackson, Mississippi.
Find out the literary antecedents for the movie, The Maltese Falcon.
Find out the tourist attractions in the eastern part of Kentucky.

Also be prepared to answer:
What is special about this engine's search techniques -what sets it apart?
Do you like or dislike this engine - why?

How Advertisers Control Search Rankings
Search engines offer all sorts of advantages over traditional ad channels, including add-on software that allows advertisers to create programs that interact with search-engine index and marketing systems. They can integrate their inventories so that ads are automatically triggered for overstocks or track ad clicks and add text to ads. Search engines auction keywords to the highest bidding advertisers. They see their ads placed alongside search results corresponding to the term and pay only when someone clicks on the ad. AOL, Ask, Microsoft, Google and Overture offers such marketing. Many are convinced search engine marketing guarantees a return on investment – making it easy to automate tasks and ads. Keyword bidding is way up. Search is now 1% of the global ad market - $5 billion.

How to Climb Search Rankings
It's no secret the Web has greatly changed how we create and retrieve information. But many used don't realize the results of their searches can be skewed by an industry called 'search-engine optimization' (SEO) that has emerged to advise people how to improve their web pages rankings in the search engines. These experts are not advising people pay for better rankings – though many major search engines or directories accept payment to increase a ranking. The large engines that do rank sites by what they pay tell users they do so – but they hide it a bit – a fact that got them in trouble with the FCC.

Many sites promote keyword buying – though this is for relevant ads. Sites can fool the search engine by placing metatags (which describe the page in the invisible portion of the page our browsers don't usually show us) that are irrelevant to their content. This is called keyword stuffing and other trick attempts will continue - search engines have gotten smarter at detecting such tricks and will ban sites they catch using them but like an arms race some site owners are always trying new strategies to bypass a search engine'd defenses causing new solutions to arise and so on.

Most advice about how to increase a site ranking is quite benign - even obvious - much of it boils down to **creating a clear and accurate title for a page** - using metatags effectively and **learning the correct submission guidelines** for each search engine.

Other advice is not so innocuous - some recommend site owners create bridge or doorway pages - written for the sole purpose of getting high rankings on search engines. - a site may have dozens of such pages, each focusing on different keywords, and each aimed at a particular engine's search ranking formula. Once you reach the bridge page, you are immediately forwarded to the real home page. This could be a detour from the topic you are searching. More troublesome is the conflation of advertising and editorial content on the web - if marketers want to get their sites listed early in search results, they need to offer more than a sales spiel, price info and product features - they need good content. The trouble for searchers is there's a great deal of difference between info on alternative energy found at Exxon's site and a scholarly article on the subject at a university's web site. Another reason highly ranked pages are becoming more commercial is because sites selling something have the greatest incentive to ensure their pages are ranked highly - thus they are most likely to invest the most resources in trying to manipulate search engines. Google seems to weed these out better. Searchers find ostensibly hard facts on all kinds of sites, and need to determine who is really behind a site and what their mission and qualifications are.

Unfortunately, it's not easy to evaluate sites as it used to be - site owners can now register .com or .org regardless whether they are commercial or nonprofit sites though edu & gov is still limited. What can you do to avoid being lured to less substantive sites? Improve your search strategies by adding more keywords to your search - which should give you fewer, but more relevant results. And use boolean algebra - "not" or the minus sign to exclude terms like 'order' and 'buy' which appear on many commercial sites or the names of companies that turn up often on less-relevant pages in a search. Although it is true that the sites at the top of rankings are most likely to be most relevant, don't neglect others on your list - especially if those atop are commercial.

You can also get better results by being careful about how and where you search. Use sites you trust. Online dbases available through your library are very useful. But don't limit your research to the web - most of the world's knowledge still resides in print.

Finally, watch out for commercial sites masquerading as informational ones - always find out who is behind a site and what their mission is - look for a link on the page to a sponsoring norganization - if you find it - note the organization's purpose - check out the other affiliations of any directors or board members listed. Look for any political ideological or commercial agenda which could detract from the site's credibility. If you can't find any info on the site itself - try a search for the site's owner on a search engine - whether the results are helpful manipulative or downright deceptive, the owner's attempts to obtain a high ranking is a force worth reckoning with.

Five Ways to Make Your Website More Search-Engine Friendly

1) Some keywords generate better quality traffic than others and may convert browsers to customers. 'adjustable' vs. 'mortgage.' Come up with narrow specific keywords, such as 'South Florida, mortgage" Figure out what search terms you want to 'own' and develop lots of high quality content around those terms. If you are authoritative in your industry, flaunt it – optimize and tag this content correctly.

2) Address your specific keyword list in your meta-tags, bearing in mind, bearing in mind that your title tag is most important. Don't stuff keywords into meta-tags. Write naturally descriptive meta-tags as well written sentences that include your keywords.

3) Get someone with a site that has a higher page rank to contextually link to you. This reinforces your site's authoritativeness on a certain search term.

4) Regularly search the name of your company and see what appears. Set up a Google Alert to be notified when your organization is mentioned.

5) Use online tools to gauge what people are saying about your company.

Tips to Increase Ranking

A high ranking in search results is like free advertising for your website. Try these tips to increasing traffic and topping the charts.

Because spiders, resulting indexes and algorithms vary by engine, it is difficult to maintain top ranking across all but here are some hints: 1) **Relevant page titles** and body copy can influence rankings - determine which keywords visitors would use to find your site - and for competitor's sites and include those terms in your pages. 2) **keyword placement** and frequency throughout a page are prime factors in the ranking process. It's a good idea to use a combination of unique and common keywords, but don't overuse unique or proprietary terms at the expense of common or general terms that are more likely to be used in a search.

Get a feel for the competition: look at other sites and their traffic

Many, but not all, search utilities also reference metatags - invisible HTML tags within documents that describe their content - as a way to control how content is indexed. As a result, proper use of metatags throughout a site can also boost search engine ranking. There are well over 50 metatags available for use within HTML pages but those most commonly used by search engines for indexing and ranking purposes are the description and keyword tags. As the name indicates, the description tag includes a short description of a site that a search engine would display in a search results list. Keywords for a search engine to associate with a particular page are stored in the keywords tag. These are words that a user would type as part of a search engine query to bring up relevant web pages from a search engine's index. The idea is to anticipate the keywords potential visitors would use to find your site. For large, content-rich sites, maintaining meta-tags can quickly become tedious. A number of products (Watchfire) will automate this process though. Of course, metatags are still ignored by some search engines, Lycos, for example ignores them because of metatag spamming - or keyword stuffing - repeating a particular keyword hundreds of times in a page's code. This practice or the use of inappropriate tags will get your site banned by many search engines.

Keeping a site high in rankings can be a fulltime job. As people arrange their websites to link to your site, and the number of these links increase, your rankings will increase.

Find hobbyist or specialty 'magazine' websites in your industry and submit press releases to them.

Homework

Email dmurphy@depaul.edu with your answer to the following questions in complete sentences. Please include your name and DePaul ID number.

Consider your work, volunteer, personal activities or home life. When could the advanced features of Word help you to improve documents that you develop, or that you use (e.g., school communication)? What are the most difficult parts of inserting a picture into a document? What tips do you have for others?
Create and email to me as an attachment a 3 page Word document titled: Creating Documents Using Microsoft Word.
The words you use are not important. Use any content that you cut and paste from anywhere you wish as ‘filler' for your edits. Please note because we are focusing on the formatting, and not the content itself, this is covered under 'fair use' and not a copyright or plagiarism issue. Please show the following in your document:

- Double-spaced lines of text.
- Centered headers for each page, with page numbers and the document title.
- Headers for each section (collection of paragraphs) of related content (flushed left).
- Boldface text.
- Italicized text.
- Three different font sizes.
- A table containing 2 columns and 3 rows.
- Two footnotes references - you can make them up, but they should be 'correctly' done.
- 1 and 1/4 inch margins.
- Your name and DePaul student ID number.

(10 points for completion)